

For Reference

NOT TO BE TAKEN FROM THIS ROOM

For Reference

NOT TO BE TAKEN FROM THIS ROOM

Ex LIBRIS
UNIVERSITATIS
ALBERTAENSIS





Digitized by the Internet Archive
in 2020 with funding from
University of Alberta Libraries

<https://archive.org/details/Ryan1969>

THE UNIVERSITY OF ALBERTA

A VALIDATION STUDY OF THE FIRO-B

BY



BRUCE ALLEN RYAN

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE
OF MASTER OF EDUCATION

DEPARTMENT OF EDUCATIONAL PSYCHOLOGY

EDMONTON, ALBERTA

FALL, 1969

THEORY OF THE EARTH

CHAPTER I. OF THE ORIGIN OF THE EARTH.

THE EARTH, as we see it, is a globe, or sphere, of a very imperfect sphericity. It is composed of a solid nucleus, or core, surrounded by a fluid envelope, or atmosphere. The solid nucleus is composed of a mass of matter, the density of which is not known, but which is supposed to be of a very great density.

CHAPTER II. OF THE ORIGIN OF THE ATMOSPHERE.

THE ATMOSPHERE, as we see it, is a fluid envelope, or atmosphere, of a very imperfect fluidity. It is composed of a mass of matter, the density of which is not known, but which is supposed to be of a very great density.

CHAPTER III. OF THE ORIGIN OF THE OCEANS.

THE OCEANS, as we see it, are a fluid envelope, or atmosphere, of a very imperfect fluidity. They are composed of a mass of matter, the density of which is not known, but which is supposed to be of a very great density.

CHAPTER IV. OF THE ORIGIN OF THE MOUNTAINS.

164(5)
208

UNIVERSITY OF ALBERTA
FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled "A Validation Study of the FIRO-B" submitted by Bruce Allen Ryan in partial fulfilment of the requirements for the degree of Master of Education.

ABSTRACT

The purpose of the present study was to evaluate the construct validity of FIRO-B, the principle measuring instrument in Schutz's (1966) theory of interpersonal behavior. Schutz postulated that every individual possesses three interpersonal needs: inclusion, control and affection. FIRO-B was designed to measure each of the needs with six 9-item Guttman scales.

The research procedure was based on Loevinger's (1957) concept of construct validity which, in this view, has three components: the substantive, the structural and the external component.

The subjects consisted of one hundred forty-four non-college adults from three distinct groups: forty-eight salesmen, forty-eight policemen and forty-eight service volunteers (twenty-four males and twenty-four females).

FIRO-B was evaluated as only partially meeting the requirements of the substantive component because the items did not appear to adequately reflect the breadth of the theory.

As Schutz chose the Guttman scale as the structural model for the subtests it was necessary to show that the scales had reproducibility coefficients of at least .90 before the requirements of the structural component could be considered met. Only the inclusion scales reached the

critical level; consequently the structural component criteria were only partially satisfied.

To evaluate the external component, three groups of subjects were chosen to serve as criteria for each of the need areas. Salesmen, policemen and volunteers were hypothesized to score high on the inclusion, express control and affection scales respectively. Other behavioral data, relevant to the needs were also collected.

In general the results from the analysis of the external component suggested that only two or perhaps three constructs, instead of six, were being measured by FIRO-B. The major dimension appeared to be predominantly inclusion-like (interpersonally participating) behavior; consequently FIRO-B seems to simply measure magnitude of interpersonal behavior. The control items, as well as reflecting the "participating" factor, appeared to relate to a control dimension of a bipolar nature.

As a measure of the basic interpersonal needs outlined in Schutz theory, FIRO-B lacks satisfactory construct validity.

ACKNOWLEDGEMENTS

The writing of this thesis would have been considerably more difficult without the assistance of several very helpful people.

I would like to express my appreciation to Dr. T. O. Maguire for offering support and guidance, particularly at the very moments when I began to despair.

To Dr. P. W. Koziey and Dr. D. Wilson I wish to express my gratitude for helpful and encouraging comments.

Much credit is due to my sister, Miss Toni Ryan, for her many hours of reading, critical comment and laborious typing.

Without the assistance of Inspector G. J. Bishop of the City of Edmonton Police Department, Mr. A. G. Stanton of the Financial Life Assurance Co., and Mrs. F. Yarrow of the Central Volunteer Bureau, the problem of data collection would have been an almost impossible task. May I express my gratitude for their invaluable help.

TABLE OF CONTENTS

CHAPTER		PAGE
I.	INTRODUCTION	1
II.	THEORETICAL ORIENTATIONS	6
	Schutz's Three Dimensional Theory of	
	Interpersonal Behavior	6
	Introduction of the Interpersonal Needs.	9
	Inclusion, Control and Affection	
	Behavior	10
	Justification for Defining the Three	
	Need Areas	12
	Development of the Interpersonal Needs .	13
	Interaction of Individuals	14
	Construct Validity	15
	Loevinger's Concept of Validity.	17
	Test Responses as Behavior	18
	Components of Construct Validity	19
	The Substantive Component	20
	The Structural Component	21
	The External Component	25
	Guttman Scalogram Analysis	27
	The Nature of a Guttman Scale.	27
	The Concept of Reproducibility	28
	The Reproducibility Coefficient.	29
	The Cornell Technique.	30

CHAPTER	PAGE
The Goodenough Technique . . .	33
III. FIRO-B AND PREVIOUS VALIDITY RESEARCH. .	37
FIRO-B	37
Previous Validity Research	40
IV. RESEARCH PROCEDURE	53
The Structural Component	53
Sample	54
Data Collection	54
Statistical Analysis	55
Hypothesis	55
Hypothesis 1	55
The External Component	55
Inclusion.	56
Control	56
Affection.	57
Samples	58
Statistical Analysis	59
Hypotheses	59
Hypothesis 2	59
Hypothesis 3	59
Hypothesis 4	59
Hypothesis 5	59
Hypothesis 6	59
Hypothesis 7	59

CHAPTER	PAGE
V. RESULTS AND DISCUSSION	60
The Substantive Component	60
The Structural Component	64
Hypothesis 1	64
The External Component	66
Hypothesis 2	67
Hypothesis 3	67
Hypothesis 4	67
Hypothesis 5	73
Hypothesis 6	73
Hypothesis 7	75
Discussion of the External Component .	75
The Construct Validity of FIRO-B . . .	80
VI. IMPLICATIONS AND SUMMARY	82
Implications	82
Summary	83
REFERENCES	88
APPENDIX A: FIRO Theory.	92
APPENDIX B: FIRO-B Questionnaire	97
APPENDIX C: Answer Sheet	100
APPENDIX D: Scale Items and Keyed Response	
Alternatives	102
APPENDIX E: Information Sheet	105
APPENDIX F: FIRO-B Scale Intercorrelations . .	107

CHAPTER	PAGE
APPENDIX G: Volunteer, Sales, and Police Group Scores on Want Control, Analysis of Variance and Scheffe Comparisons Among Means	109
APPENDIX H: Scale Means and Scheffe Comparisons Among Means on the Six FIRO-B Scales for Female Volunteers, Male Volunteers, Salesmen, and Policemen	111

LIST OF TABLES

TABLE		PAGE
1.	Summary of Parent-Child Relations and the Resultant Adult Behavior	14
2.	The Cornell Technique	32
3.	The Goodenough Technique.	35
4.	Scores on FIRO-4 Various Occupational Groups.	48
5.	Reproducibility Coefficients for Each of the Six FIRO-B Scales	65
6.	Unrotated Principal Axes Factors Based on Scales Scores of One Hundred Forty-Four Subjects	65
7.	Mean Scores and Standard Deviations for the Volunteer, Sales, and Police Groups on the FIRO-B Scales	66
8.	Express Inclusion Group Means, Analysis of Variance, and Scheffe Comparisons Among Means	68
9.	Want Inclusion Group Means, Analysis of Variance, and Scheffe Comparisons Among Means	69
10.	Express Control Group Means, Analysis of Variance, and Scheffe Comparisons Among Means	70

TABLE		PAGE
11.	Express Affection Group Means, Analysis of Variance, and Scheffe Comparisons Among Means	71
12.	Want Affection Group Means, Analysis of Variance, and Scheffe Comparisons Among Means	72
13.	Club and Non-Club Group Means on Express Inclusion and Summary Analysis of Variance	74
14.	Executive and Non-Executive Group Means on Express Control and Summary Analysis of Variance	74
15.	Executive and Non-Executive Group Means on Want Control and Summary Analysis of Variance	75
16.	Express Control Mean Scores for Male and Female Volunteers, Salesmen and Policemen: Analysis of Variance and Scheffe Comparisons Among Means.	78

CHAPTER I

INTRODUCTION

To a very large extent, the usefulness of a personality test is determined by its validity. If a test is to remain a productive research and diagnostic instrument, its validity must be reaffirmed in an ever-widening circle of circumstances. Following this dictum the present study is concerned with an examination of the validity of FIRO-B (Fundamental Interpersonal Relations Orientation - Behavior level).

FIRO-B, a 54-item questionnaire containing six subtests, was developed by Schutz (1958) to measure three important personality characteristics: inclusion, control, and affection. They were defined by Schutz (1958) as follows:

Inclusion: ... the need to establish and maintain a satisfactory relation with people with respect to interaction and association [p. 18].

Control: ... the need to establish and maintain a satisfactory relation with people with respect to control and power [p. 18].

Affection: ... the need to establish and maintain a satisfactory relation with others with respect to love and affection [p. 20].

These three variables are defined and interrelated within Schutz's (1958) three dimensional theory of interpersonal behavior.

A cross-validation of FIRO-B seems indicated at the

present time for a variety of reasons. The first stems from a general scientific standpoint. Campbell and Fiske (1959) stated that "before one can test the relationships between a specific trait and other traits, one must have some confidence in one's measures of that trait." Any findings that strengthen that confidence increase the value of the test and the theory underlying the test.

A second source of impetus for further validation rests in two important features of the development of FIRO-B. The first relates to the gradual evolution of Schutz's theory and, concomitantly, the test, which is the pivotal measuring instrument for the theory. During this process, validity data did become available and much of this information is reported in Schutz's 1958 book. However, all of the studies cited involved earlier versions of FIRO-B such as FIRO-Q, FIRO-1, FIRO-2, FIRO-3, FIRO-4, and FIRO-5B3. While these led to findings supporting the theory in general, they did not supply directly validating evidence for FIRO-B itself. While the present study is concerned with the gathering of data bearing directly on FIRO-B, methods similar to some utilized by Schutz were applied.

The second important feature of FIRO-B's development relating directly to its validation lies in the hypothesized structural quality of the test items. Because Schutz was primarily concerned with the measurement of specific

orientations, the Guttman technique for cumulative scale analysis was chosen as the most appropriate model on which to base the test. Each of the six subtests was constructed in accordance with the principles governing Guttman scales. Several writers, including Loevinger (1957), and Tukey (1950) have emphasized that a cross-validation procedure is critical when such a formal structure of items is proposed. In fact, with the relatively undefined variables in personality research this procedure is of double importance. Tukey (1950) stated:

In every field of science, and particularly in fields where data and analysis [are] complex, there are two different phases of quantitative analysis -- exploration and confirmation -- and almost always, when dealing with complex problems, these have to be carried out on different samples of data [p. 66].

Schutz did carry out such a cross-validation procedure and found that the scale characteristics of his items were maintained on the second administration. Almost all the subjects used in the original standardization procedure and cross-validation were college students. The purpose of the present study, in part, was to extend the cross-validation for the scale characteristics to a non-college adult population. Because the test has been designed to apply to all people it is important that subjects other than college students be used for validation purposes. One of the most frequently used techniques for establishing a test's

validity is to determine how well it distinguishes among subjects possessing varying amounts of the traits being measured. Schutz has already obtained data of this nature on earlier forms of FIRO-B and these are discussed in a subsequent chapter. One of Schutz's validation methods was to compare the score patterns for various occupational groups. The present study utilized a similar procedure but with different occupational groups.

The theoretical model for the validation procedure followed was that outlined by Loevinger (1958). While her's is basically a "theory of the first test," her views have important implications for any validation study. To Loevinger it is critical that test and theory be validated together and if a test is to contribute to scientific growth in psychology, it must have some systematic and meaningful relation to theory. Simple correlational and mechanical demonstrations of the relationships among test scores and external criteria, as exemplified by the procedure of pure empirical keying, are not adequate as a theory of psychometrics. These methods contribute only minimally to the understanding of how test items relate to behavior theory. Consequently, in Loevinger's view, construct validity is the only viable object of study in terms of psychological theory because of its emphasis on the relation of the test item to the psychological "thing" or trait being measured.

.

Construct validity, as presented by Loevinger, has three main components: a substantive component, a structural component, and an external component. Each of these concepts will be defined and discussed in Chapter II. For a test to have any real psychological validity, the requirements of each component must be met; otherwise no truly meaningful interpretation of scores can be made in the context of the theory underlying the test.

The essential purpose of the present study was to collect evidence as to the validity of FIRO-B for a non-college population using Loevinger's conception of construct validity as the basis for the research procedures and data analysis.

CHAPTER II

THEORETICAL ORIENTATIONS

In the present chapter, a description and a delineation of the theoretical problems and issues involved in the study is presented. The first section deals briefly with Schutz's (1966) interpersonal theory, and an attempt is made to demonstrate the very intimate relationship between FIRO-B and the theory.

The next section is devoted to a summary and analysis of Loevinger's (1957) theory of psychometrics stemming from the concept of construct validity. Her work served as the theoretical basis for the validation procedure adopted.

Thirdly, the concept of scalogram analysis is developed in some detail. FIRO-B was constructed as a Guttman scale and a cross-validation of the test's scale characteristics is an important aspect of the validity analysis followed. Considerable attention is devoted to the idea of reproducibility and two of the main methods for calculating a reproducibility coefficient are discussed.

I. SCHUTZ'S THREE DIMENSIONAL THEORY OF INTERPERSONAL BEHAVIOR

FIRO, Fundamental Interpersonal Relations Orientation (see Appendix A), was first formally described as a general theory of interpersonal behavior by Schutz in 1958 in FIRO:

A Three Dimensional Theory of Interpersonal Behavior. In 1966 a reprint edition appeared and was renamed The Interpersonal Underworld.

Aside from the actual content, Schutz's theory has two rather interesting features, both of which are concerned with the general problem of theory construction. Schutz has created his theory in the form of the hypothetical-deductive model with a formal system of postulates and theorems which he felt are readily testable. As Borko (1960) observed, he deserves credit for even making the attempt.

In conjunction with the formal organization, Schutz has followed the rule of Bayes for presenting, describing, and arguing the case for the truth of his postulates. He chose this approach because he felt the Bayesian model to be "the most satisfactory formulation of the method of relating evidence to hypothesis ... [Schutz, 1966, p. 8]."

Briefly and superficially stated, the rule of Bayes asserts that the truth of a statement is the function of two factors. The first, the antecedent probability, is the degree of belief that a statement is true when all the previously acquired knowledge is evaluated. The second, the evidential probability, refers to the probability that a particular type of evidence would occur if the proposition were true. This latter sort of information is the kind supplied by the classical experimental paradigm.

Rather than being a detailed application of the concepts of Bayesian statistics, Schutz's presentation only used the Bayesian approach as an organizing model. For each postulate in the system he carefully specified what other authors have written on the topic (antecedent probability) and has supplied fairly detailed discussions of his own experiments (evidential probability).

At the base of this structure of postulates lies a general and most germane assumption which Schutz accepted as a given. The entire theory, as Schutz stated, "may be considered an elaboration of the phrase 'people need people'."

In Schutz's terms, people need people in three distinct ways, referred to as needs which can be thought of as having properties very similar to biological needs. The three needs, inclusion, control, and affection, were previously defined. A further elaboration of their nature will be presented here.

Each need has two important aspects: people need to receive from others and need to express toward others. Consequently a total of six separate variables become important in any attempt to understand a given instance of interpersonal interaction.

Also, an integral part of the nature of the interpersonal needs is that all can be thought of as covering a parameter of observability ranging from not observable to

anyone but the subject himself to occupying a public domain observable to all. Schutz considered this parameter as being dichotomized into a behavioral domain (external and public) and a feeling domain (observable only to the subject himself). The latter area, while an aspect of FIRO, is less fully developed at the present time and is beyond the scope of the present study. FIRO-B is a measure of the needs in a public domain.

The typology is organized around four major postulates which will be briefly discussed below. For the purposes of this study, the first, which introduces the interpersonal needs, will be described in the greatest detail. The others, which are concerned with the development of the needs and the interaction of individuals, will receive only a cursory examination.

Introduction of the Interpersonal Needs

Postulate 1. The Postulate of Interpersonal Needs.

(a) Every individual has three interpersonal needs: inclusion, control, and affection.

(b) Inclusion, control, and affection constitute a sufficient set of areas of interpersonal behavior for the prediction and explanation of interpersonal phenomena (Schutz, 1966, p. 13).

It is clear that Schutz believed that every individual has each of the three needs and that these needs account for what takes place in an interpersonal situation.

An interpersonal situation is one involving two or more people such that at least one has an influence on the

other's behavior. Two people in a discussion constitute an interpersonal situation. A man, seated on a bus, debating with himself as to whether or not he should give his seat to a standing female passenger, is involved in an interpersonal situation even though she may not even be aware that he exists. For her the situation need not be an interpersonal one.

As mentioned previously, the interpersonal needs are conceptualized as having properties closely parallel to biological needs. A biological need requires the establishment of a satisfactory relation between an organism and its physical environment. An interpersonal need requires the establishment of a satisfactory relation between the individual and his human environment. An "optimal interpersonal relation" is required to bring about satisfaction in the interpersonal sphere. Any departure from the optimal conditions results in anxiety.

Inclusion, control and affection behavior. Each of the three needs has associated with it particular behaviors that are directed toward the satisfaction of that need. Below is an expanded description of the behaviors stemming from each need.

"Association with people" is the province of inclusion behavior. Some terms that describe inclusion behavior are: interact, mingle, attend to, member, join, exclusion, isolate,

lonely, abandoned, and ignored. Attention-seeking behaviors are primarily inclusion based. Fame is inclusion behavior in the extreme. Just to be known by a great many people is satisfying to the high need for inclusion type. People who join clubs may be considered inclusion types.

Control behavior is concerned with power, and decision making in interpersonal matters. Some terms describing this relation are: power, authority, dominance, superior officer, leader, rebellion, follower, anarchy, submissive, milquetoast. Political power and the acquisition of economic resources are often control behaviors. The control seeker argues to win while the inclusion seeker simply enjoys the interaction.

"There is no necessary relation between an individual's behavior toward controlling others, and his behavior toward being controlled [Schutz, 1958, p. 24]." This differs from the other need behaviors where there is a tendency to express and want the same behaviors. In the case of control, for example, the rebellious adolescent may react strongly to any control being placed on him, yet be quite happy to be the leader in the neighborhood gang.

Affection behavior refers to close personal emotional feelings between people. Schutz stressed that this relationship is a dyadic one and can only occur between two people at a time. Terms denoting this need are: love, like, emotionally close, personal, sweetheart, hate, dislike,

and, cool. An affection relationship involves very close ties with others and a sharing of intimate feelings and anxieties. In a group a person can develop friendships with specific members (affection behaviors) or he may become popular and treat all in an equally friendly manner. This latter behavior is more likely to be inclusion behavior.

Justification for defining the three need areas.

Schutz was able to supply some support for his contention that these three needs explain all important interpersonal behavior. Following a Bayesian approach, he began arguing for the plausibility (antecedent probability) of his postulate by citing what other writers have hypothesized. The sweep of his discussion included Freud (1950), Fromm (1947), Horney (1945), Champney (1941), Leary (1958), and Benne and Sheats (1948) among others. The core of his conclusions with regard to these authors is that their theories have included some aspects of the three interpersonal needs. In effect, they have been describing the same elephant in its various aspects. Schutz felt that he succeeded in describing the whole elephant and that he actually incorporated the others' ideas into his own position. As Douglas (1961) noted, "the terrible audacity of the work is captivating."

Unfortunately, no really defensible experimental evidence (evidential probability) is provided to help establish the truth of Postulate 1. A single factor

analytic study was included but, as Schutz (1966) himself noted in his preface, its support is extremely weak and probably should be omitted from the text.

Development of the Interpersonal Needs

Postulate 2. The Postulate of relational Continuity.

An individual's expressed interpersonal behavior will be similar to the behavior he experienced in his earliest interpersonal relations, usually with his parents, in the following way:

Principal of Constancy: When he perceives his adult position in an interpersonal situation to be similar to his own position in his parent-child relation, his adult behavior positively covaries with his childhood behavior toward his parents (or significant others):

Principle of Identification: When he perceives his adult position in an interpersonal situation to be similar to his parents' position in his parent-child relation, his adult behavior positively covaries with the behavior of his parents (or significant others) toward him when he was a child (Schutz, 1966, p. 81).

Schutz recognized that the central concepts involved in Postulate 2 are not new ideas but that the position is "... essentially a restatement of a fundamental aspect of psychoanalytic and other psychodynamic theory." The family setting becomes the training ground for interpersonal skills. People, at a fairly early age, learn their characteristic ways of interacting and, depending on the pattern of rewards and punishments, the strengths of the interpersonal needs become established and persist into adult life.

Table 1 provides a description of the kinds of behavior that may result from particular parent-child relations.

TABLE 1
SUMMARY OF PARENT-CHILD RELATIONS AND THE RESULTANT
ADULT BEHAVIOR

	This type of parental behavior toward child	is called	and leads to this type of behavior by the child through to adulthood
Too Much	Inclusion	Enmesh	Oversocial
	Control	Dominate	Autocratic
	Affection	Smother	Over-personal
Too Little	Inclusion	Abandon	Undersocial
	Control	Undirect	Abdicrat
	Affection	Unlove	Underpersonal
Ideal	Inclusion	Include	Social
	Control	Respect	Democrat
	Affection	Love	Personal

Note: Reprinted from W.C.Schutz, The interpersonal under-world. Palo Alto: Science and Behavior Books, 1966.

Interaction of Individuals

The second half of Schutz's (1966) book, containing Postulates 3 and 4 (see Appendix A), is devoted to an examination of the principles of the interpersonal theory in their relevance to the study of group behavior. Postulate 3 is concerned with compatibility of the group members and how this variable relates to group functioning and the group's ability to move toward a goal. Compatibility is discussed in terms of the need areas in both their "want"

and "express" aspects. Various combinations of these six variables, defined by FIRO-B in all cases, lead to various kinds of compatibilities and incompatibilities.

Clearly, the role of FIRO-B in this area is critical to the contributions Schutz's theory can make. If FIRO-B is valid, the theory can move into a study of the problems of compatibility. If it is not valid, the theory becomes functionally without value because no other means is available for measuring the needs and thus defining the indices of compatibility.

In Postulate 4, Schutz outlined his hypothesis of how a group develops from a collection of strangers to a smoothly functioning and interacting group of people.

The role of FIRO-B in practical applications for this theoretical structure is extremely important. Although the instrument is not critical to the theorizing and discussion at an abstract level, experimental investigation is completely dependent on FIRO-B. No other test exists that provides a measure of the variables as they are conceptualized and defined within the typology. In effect, any operationally based experimental work on the theory is impossible without data from FIRO-B. The question of the test's validity is of the highest priority.

II. CONSTRUCT VALIDITY

Validity, globally defined, is concerned with the

extent to which a test measures what it is supposed to measure. This definition, although logically consistent with all current interpretations of validity, is sufficiently vague to generate disagreement about what the idea actually means in practice. One general position, referred to in the literature as the "classical position", sees validity residing in how well a test correlates with a particular, external behavioral event. In this sense a test has several validities and is, in a manner of speaking, capable of measuring more than one "thing". Gulliksen (1950) noted that "... validity cannot be regarded as a fixed or a unitary characteristic of a test (p. 88)." In reality there are as many validities as there are uses for a test.

A second general position, historically more recent than the first, views tests as measuring characteristics and traits that exist in people. The main purpose of a test, from this viewpoint, is to provide information on psychological constructs. This kind of validity has been called construct validity (Cronbach and Meehl, 1955; Standards for Educational and Psychological Tests and Manuals, 1966). While the specific experimental procedures applied to establish construct validity do not differ substantially from the more mechanical operations of classical validity, the stress is placed on interpretation of research findings. An effort is made to uncover concepts or constructs which

account for test performance. The application of rational analysis is as important as a rigorous pursuit of sound experimental technique. Validity of this type is more resistant to administrative changes and contextual shifts. The result is a concept of validity that is relatively unitary and stable in contrast to the highly fragmented interpretation of the classical position.

Loevinger's Concept of Validity

Loevinger (1957) argued that any theory of psychometrics must have as its central concept the idea of construct validity. If tests are to be useful instruments of psychological science they must relate meaningfully to the underlying constructs involved in determining an external behavioral event.

Consequently, Loevinger maintained that "classical validity" is not a suitable basic concept for test theory; it does not provide an adequate basis for test construction. Her outright rejection of the classical approach places her in a position of sharp disagreement with Cronbach and Meehl (1955) and the official position of the American Psychological Association. Cronbach and Meehl, whose stance on this point reflected that held by the APA, stated that they did not see construct validity as being preferable to or as replacing the other three types (concurrent, predictive, content). They did, however, indicate that construct validity should be given more attention in research than had

been the case.

Loevinger, on the other hand, did not hesitate to elevate construct validity to a position of superiority and suggested that the other concepts of validity were non-essential. In fact, she stated that perhaps less confusion might have resulted in the literature if construct validity had been given a name such as "essential validity".

Construct validity must be placed in this superordinate position because it is the only "... kind of validity which exhibits the property of transposability or invariance under changes in administrative setting which is the touchstone of scientific usefulness [Loevinger, 1957, p. 641]."

Test responses as behavior. A fundamental assumption being made by Loevinger in her analysis of the nature of validity concerns the relationship between test responses and behavior in general. A test response may be interpreted as a sign indicating the presence of some trait. Such a consideration allows the psychologist to infer from the test response itself something about the broader aspects of a particular trait. However, a test theory limited to this conception becomes dominated by mechanical sets of operations using the correlation idea as its central level. For example, it may be found that the majority of successful bankers wear brown shoes. Brown shoes then become a sign for success in the banking industry and the user of this "test for banking" would tend to make the correct

decision about a client's banking skills simply by noting the color of his shoes. Loevinger decries this procedure as lacking scientific strength and points out that test responses are also samples of behavior.

As samples of behavior, test responses can "... be assumed to obey the same laws and be determined by the same factors as other behavior [Loevinger, 1957, p. 644]." As a behavior sample, a test response can directly reveal the activity of the trait under consideration and contribute significantly to the confidence of an interpretation of that response. A question that relates to the essential qualities of banking will allow the banking trait in each subject to be evaluated directly and does not depend on a fortuitous correlation between the trait and some non-essential bit of behavior, such as the wearing of brown shoes. Test responses must be viewed as samples of behavior before the relationship between a test and the underlying psychological construct can be understood.

Test responses, then, are at once samples and signs of behavior in general. As samples they reveal the laws governing behavior and as signs they permit inferences about other behavior not sampled by the test. The ability of a test to provide this kind of information is demonstrated by its construct validity.

Components of construct validity. Loevinger stated that there are three components of construct validity that

are mutually exclusive, exhaustive and mandatory. The three relate directly to the main steps in test construction whereby the first step is to create a pool of potential test items. Next, the internal structure of this pool is analyzed and, on the basis of this information, items are selected to form the test. A final step involves obtaining correlations of test scores with external criteria. The components have been called: the substantive component, the structural component and the external component.

(1) The Substantive Component.

The substantive component of validity is the extent to which the content of the items included in (and excluded from?) the test can be accounted for in terms of the trait believed to be measured and the context of measurement [Loevinger, 1957, p. 661].

The concept of the substantive component is somewhat akin to the classical idea of content validity but the former is slightly broader in its terms of reference. To establish substantive validity it is not enough to simply demonstrate that an item represents and is part of the domain under study. Rather the inclusion of an item in the test must be explainable and justifiable from what is known about the characteristic being measured. A specific test item must also be understood in its relation to the rest of the items in the test.

Considerations of the logical relations between test items and theory alone cannot serve as an adequate basis for

test construction. Empirical evidence must also be obtained, this being the province of the last two components of construct validity. Once the pool of potential items is obtained, inter-item relations and correlations with external criteria can serve as guidelines for the final choice of items. But in the end there must also be a logical explanation for the presence of every item in the test. Such an approach avoids the folly of relying solely on pure empirical keying (e.g., brown shoes equal banking) and, at the other extreme, the lack of information on the structural characteristics of a test that would result from a consideration of content alone.

When questions concerning the substantive validity of an already established test, such as FIRO-B, are being asked, it is necessary to consult the descriptions of the constructs under investigation. If the test items appear to reflect the range of the constructs and if a logical argument can be put forth to relate items and theory, then evidence exists that the test has some degree of substantive validity.

(2) The structural component.

The structural component of validity refers to the extent to which structural relations between test items parallel the structural relations of other manifestations of the trait being measured [Loevinger, 1957, p. 661].

The task of the test builder, as well as defining a

pool of items, is to establish what structural model best describes the trait he is interested in measuring.

Loevinger noted that there are probably many possible structures existing in psychological characteristics. Most of these, however, are included under three general types: quantitative models, class models and dynamic models.

Under a quantitative model, the amount of a trait possessed is directly proportional to the number of its manifestations. The class model, on the other hand, views manifestations of a trait as simply an indication that the trait exists, making no attempt to measure quantity. The dynamic model is intended to cover those cases where opposing behavioral tendencies can represent a common source, i.e., a trait.

The characteristics of the quantitative model will be discussed here in greater detail because of its special relevance to FIRO-B which is based on a particular type of quantitative model, the Guttman scale. The Guttman scale itself will be described in detail in a later section of this chapter.

As stated above, the quantitative model is concerned with measuring the amount of a trait and uses instances of trait manifestation as its indicator of quantity. The basic assumption made in this case is that the characteristics in question are distributed along a continuum from a lesser to a greater amount, and that by asking a sufficiently

wide variety of questions, representative of that continuum, an investigator can distinguish among individuals who differ in the amount of the trait they possess. Different methods are available for selecting test items that can accomplish this task.

One method is called the "classical quantitative model" by Loevinger and it simply employs a homogeneous set of items of approximately the same difficulty or its analogue. An individual's accumulated score on the set becomes a direct indicator of the amount the trait he possesses. No further assumptions are made concerning the interrelationships among the items. No one item is considered more "important" than any other item. Under this system, however, it is possible for a number of people to obtain identical scores even though they pass and fail entirely different items in a given set. Having obtained identical scores they are then considered identical with respect to the trait yet are in complete disagreement with respect to the items.

A second method also involves the use of highly homogeneous items but spaces them on a difficulty continuum. The result is that some items carry heavier weights than others and will not be scored on unless a substantial amount of the trait is present. Consequently, individuals with the same score have likely obtained it from the same items. The Guttman scale employs this type of approach.

The actual interrelationship among a set of items is only part of the concept of structural validity. A collection of items may be shown to have a mathematically attractive inter-item arrangement and to rigorously follow a mathematical model, but this does not necessarily mean that the items are a valid measure of any specific psychological trait. It must also be shown that the model corresponds to the structure of the trait's non-test manifestations. In other words, if the test items are samples of behavior they must reflect the structure inherent in that behavior.

A model should be chosen because it accurately reflects the structure of behavior and not simply because of its mathematical rigor. Structure should be discovered rather than imposed. Loevinger pointed out that this is often a problem with Guttman scalogram analysis, where, in an attempt to obtain the .90 reproducibility coefficient, investigators may be inclined to maximize scaleability by manipulating scoring criteria, especially where multiple-choice items are involved. While the fidelity of the items to the model may be improved, the actual validity with respect to the trait under consideration may be diminished.

Once a test author has chosen a model and constructed a set of items that are true to the model, it is important to demonstrate that this fidelity is maintained under a wide set of conditions. The very important issue of cross-

validation is particularly relevant for the structural component of construct validity. The original item organization is based on the quantitative data obtained from the standardization sample and, of course, reflects the initial sampling idiosyncracies as well as the characteristics of the trait content from which the items were selected. To ensure that the items did not assume their structure due to peculiarities in the original sample, additional subjects must be tested. If the items assume the expected structure when scored by the same criteria laid down from the first analysis, the test builder has some assurance that he has fulfilled some of the requirements of the structural component.

(3) The External Component

Thus far the discussion has been directed to the correspondence between test and theory with little regard for actual behavioral expressions aside from test responses. To establish construct validity it is not enough to be able to demonstrate a logical relation between items and theory, or to show that item interrelationships adequately follow an appropriate model. Eventually the test builder must supply evidence that the test, as a measure of a trait, provides data that are consistent with non-test behavior. This constitutes the process of establishing the external component of construct validity.

In the Standards for Educational and Psychological Tests and Manuals (1966), the external component was covered

under the heading of "criterion related validity". This type of validity, the Standards states, "... is demonstrated by comparing test scores with one or more external criteria considered to provide a direct measure of the characteristic or behavior in question [p. 13]." The main result of such a procedure is to establish an interrelated network of relationships among the many criteria so that the nature and limitations of a test as a measure of a construct or trait can become well-defined. When the test author has shown that test scores do correlate with other manifestations of the trait in question, he has met the requirements of the external component of construct validity.

Construct validity, analyzed into its three component parts, is taken by Loevinger to be the only meaningful basis for a theory of psychometrics. Problems such as reliability no longer exist as relevant issues if it can be demonstrated that a test has true construct validity. Such validity cannot be claimed as established until each component has been dealt with in detail. A weakness in any one of them means that adequate construct validity has not been attained although this need not preclude the possibility of a purely mechanical application of the test in highly specialized circumstances. The test's usefulness in psychological research, however, would likely be minimal.

In the present investigation, Loevinger's conception of construct validity was applied to FIRO-B. This was seen

as especially important because of the test's intimate relationship with the theory in general, and more specifically, with the interpersonal dimensions.

III. GUTTMAN SCALOGRAM ANALYSIS

It was pointed out in a previous section that a major problem facing a test builder is to choose a model that reflects the important aspects of the trait he wishes to measure. The model, by providing a simplified version of the behavior being sampled, acts as an objective standard by which the organization of the test may be judged. Once the author has chosen an appropriate model he then must select items that fit his model. This section is devoted to a discussion of the scalogram analysis technique which was developed to select items fitting a particular model -- a Guttman scale.

The Nature of a Guttman Scale

Guttman (1950) stated that a Guttman scale exists "... if it is possible to derive from the distribution a quantitative variable with which to characterize the objects such that each attribute in the universe is a simple function of that quantitative variable [p. 88]."

Edwards (1957) phrased the definition as follows:

If a set of statements with a common content is to constitute a Guttman scale, then an individual with a higher rank (or score) than another individual on the same set of statements must also rank just as high or higher on every statement in the set as the other individual [p. 172].

A Guttman scale, then, must be able to distribute people along a continuum consisting of a single dimension. In other words, it must form a unidimensional scale. In such a scale, all items are considered homogeneous with respect to a particular dimension or trait and none relate to any other trait. Each statement has a singular meaning.

In spite of the obvious advantages of such item sets, most sociological and psychological research instruments do not form unidimensional scales with the result that scores obtained from these tests represent a mixture of variables. Unidimensional scales on the other hand, reflect a single variable. While no confusion is associated with the meaning of a particular score on a Guttman scale, no clear interpretation can be made about a score obtained from items representing several different traits.

How does an investigator know when he has found a scaleable set of items? As Festinger (1947) pointed out, he must have some means to evaluate "how well ... his data fit those conditions which would exist if unidimensionality were present [p. 155]."

The Concept of Reproducibility

Although several criteria for judging unidimensionality were stressed by Guttman (1950), the concept of reproducibility is probably the central and most meaningful one. The coefficient of reproducibility provides a measure of the amount a particular set of items deviates from the ideal

scale pattern (Guttman, 1950, p. 77).

If a scale possesses perfect reproducibility, the pattern of a subject's responses should be reproducible from knowledge of his scale score alone. This feature arises from the stepwise relationship among the items where they are ordered on a continuum of difficulty or probability of endorsement. As an example, consider a set of nine items increasing in difficulty from 1 to 9 such that fewer and fewer people correctly answer the questions as they approach number 9. For a particular individual, correct responses may be given for questions 1 to 5, with item 6 being incorrectly answered. If a perfectly reproducible scale exists in those items he will also respond incorrectly to items 7, 8, and 9. His final score on the set becomes 5, this being obtained on the first 5 items. Thus his scale score becomes a perfect summary for his actual score pattern. However, it should be stressed again that these relations hold only when the condition of perfect reproducibility has been attained.

The reproducibility coefficient. In psychological research the situation of perfect reproducibility is almost never encountered although reasonable approximations to it do occur with some frequency. A measure of the degree to which a set of items reflects this characteristic is given by the reproducibility coefficient.

This coefficient can assume values ranging from 0 to 1 depending on the frequency of errors made in recreating response patterns from total scores. The coefficient of 1 occurs in the case of perfect reproducibility. Guttman (1950) suggests that coefficients of .90 or higher sufficiently approximate the perfect case.

Of the several methods available for calculating reproducibility coefficients, only two will be discussed. A fairly complete description of the various techniques was given by White and Saltz (1957). The first to be dealt with here is the Cornell technique originally described by Guttman (1950). The second is the Goodenough (1944) procedure which utilizes a slightly different approach to error definition.

(1) The Cornell Technique. To simplify the discussion, only the case of dichotomous items will be outlined. This relates meaningfully to FIRO-B, in which each item, although presented in multiple-choice form, was dichotomized according to a key. All responses are assigned either 1 or 0. The value 1 is assigned to a response indicating a greater amount of whatever is being measured. These scores are then added to obtain the subject's scale score.

Edwards (1957) provided a clear and quite complete outline of the Cornell procedure for calculating the reproducibility coefficient from dichotomous data. The

steps outlined by Edwards are as follows:

(i) Once the items have been scored and the scale scores for each subject have been obtained, the subjects are ranked with the highest score placed at the top. The nature of their responses (1 or 0) for each statement is indicated in the same table. Table 2 provides an example of the arrangement for 20 subjects and 4 items.

(ii) The next step is to establish cutting points to make the point at which responses for each item shift from 1's to 0's (accept to reject). Two rules act as guidelines for the positioning of cutting points: error should be minimized and "no category should have more error than non-error [Edwards, 1957, p. 187]." All scores above the cutting point should be 1's and all below should be 0's. All 0's above the line are errors and all 1's below the line are errors.

(iii) All errors in each statement are then counted and added together to form the total number of errors in the array.

(iv) This total is then substituted in the following formula:

$$\text{Coefficient of Reproducibility} = 1 - \frac{\text{No. of errors}}{\text{No. of items} \times \text{No. of S's}}$$

The resulting number is the reproducibility coefficient.

Its specific purpose is "... to indicate the percent accuracy

TABLE 2
THE CORNELL TECHNIQUE

The Cornell technique applied to a 4 statement scale responded to by 20 subjects. The horizontal lines in the body of the table are possible cutting points for the statements.

Subjects	Statements								Scores
	1		2		3		4		
	1	0	1	0	1	0	1	0	
1	x		x		x		x		4
2	x			x	x		x		3
3	x		x			x	x		3
4	x		x			x	x		3
5	x			x	x		x		3
6	x			x	x		x		3
7	x			x	x		x		3
8	x			x	x		x		3
9	x			x		x	x		2
10		x	x			x	x		2
11	x			x		x	x		2
12	x			x		x	x		2
13		x	x			x	x		2
14		x		x	x		x		2
15		x	x			x		x	1
16		x		x		x	x		1
17	x			x		x		x	1
18		x		x	x			x	1
19		x		x		x	x		1
20		x		x		x		x	0
f	12	8	6	14	8	12	16	4	
p and q	.6	.4	.3	.7	.4	.6	.8	.2	
e	1	1	3	1	2	2	2	0	$\Sigma e = 12$

Note: Reprinted from A. L. Edwards, Techniques of attitude scale construction. New York: Appleton-Century-Crofts, 1957. p. 183.

with which responses to the various statements can be reproduced from total scores [Edwards, 1957, p. 183]."

The reproducibility coefficient for the data in Table 2 is: $1 - \frac{12}{80} = .85$.

Although indicating the accuracy of reproducibility may be the objective of the Cornell technique, it may be shown that it does not actually represent the real accuracy involved in reproducing score patterns from total scores (Edwards, 1948).

(2) The Goodenough Technique. Edwards (1948) favored the use of the Goodenough technique because it reflects more accurately the degree of error involved in reproducing the response patterns from the total scores. The Cornell and Goodenough techniques differ in two basic ways: in the method of determining cutting points or their equivalent and in the means of defining an error.

In the Cornell technique the cutting point was placed rather arbitrarily so that error was minimized. Each statement was considered individually and no reference was made to the pattern of scores that should be reproducible from a specific total score. Cutting points in the Goodenough method, however, are not as directly responsible for deciding whether or not a particular response is in error. Instead, the cutting points, which are consistently defined by the proportions obtaining 1's and 0's for each statement,

serve as a means of ranking the statements on a continuum of difficulty or its analogue. The statements would then be ordered in terms of their frequency of endorsement.

Table 3, using the same data as Table 2, shows the arrangement of the responses of 20 subjects on 4 items. The proportions obtaining 1's and 0's are given at the bottom of the table and the cutting points are given by the lines in various positions in the four columns. From this table, it is clear that most people scored on statement 4, then on 1, 3, and 2 in that order. Once this ranking of items has been obtained, the investigator is in a position to reproduce score patterns that should be associated with particular scale scores, or scale types, as they are often called.

For the data in Table 3, five response patterns should exist, if the items were perfectly reproducible, for the five scale types. If a subject obtains a score of 4 (scale type 4), he must necessarily receive 1 for each item. In the case of a score of 3, the expectation would be that he score 1 on items 4, 1 and 3 and 0 on item 2. According to the cutting points, item 2 should be the first item to be rejected when all cannot be accepted. For a score of 2, 1's should be obtained on items 4 and 1, but 0's are expected on 3 and 2. When a score of 1 occurs there should be a score of 1 on statement 4 and 0 on the rest. Finally, in the case of a score of 0, all items should be scored 0.

TABLE 3

THE GOODENOUGH TECHNIQUE

The Goodenough technique applied to a 4 statement scale responded to by 20 subjects.

Subjects	Statements				Scores	e
	1	2	3	4		
1	1	1	1	1	4	0
2	1	0	1	1	3	0
3	1	1	0	1	3	2
4	1	1	0	1	3	2
5	1	0	1	1	3	0
6	1	0	1	1	3	0
7	1	0	1	1	3	0
8	1	0	1	1	3	0
9	1	0	0	1	2	0
10	0	1	0	1	2	2
11	1	0	0	1	2	0
12	1	0	0	1	2	0
13	0	1	0	1	2	2
14	0	0	1	1	2	2
15	0	1	0	0	1	2
16	0	0	0	1	1	0
17	1	0	0	0	1	2
18	0	0	1	0	1	2
19	0	0	0	1	1	0
20	0	0	0	0	0	0
f	12	6	8	16	42	16
p	.6	.3	.4	.8		
q	.4	.7	.6	.2		

Note: Reprinted from A.L. Edwards, Techniques of attitude scale construction. New York: Appleton-Century-Crofts, 1957. p. 185.

Any deviations from these expectations is an error in reproducibility. These errors for Table 3 are summarized in the extreme right-hand column.

The errors are then totaled and substituted in the formula for the reproducibility coefficient given above. The reproducibility coefficient for the data in Table 3 is:

$$1 - \frac{16}{80} = .80.$$

In contrast to the Cornell technique, the Goodenough approach more meaningfully attempts to measure the accuracy of reproducing responses from knowledge of total scores. In general the numerical outcome of the more rigorous method is an increased error factor. For the above example, the number of errors increased from 12, for the Cornell technique, to 16 for the Goodenough procedure. The reproducibility coefficient decreased from .85 for the Cornell technique to .80 for the Goodenough method. In actual practice where larger samples are used, the differences between the two approaches are likely to be much greater.

Because of the more meaningful and rigorous standards associated with the Goodenough technique it was adopted for application in the present study. Schutz (1966) was not clear as to the method used in the original analysis of FIRO-B, but as Guttman (1950) was his main reference for the scalogram analysis, it is likely the Cornell technique was applied.

CHAPTER III

FIRO-B AND PREVIOUS VALIDITY RESEARCH

The present chapter consists of two general sections, both concerned with a clarification of the nature of FIRO-B as a test. The first section concentrates on a discussion of the test's development and a description of its published format. The second section is devoted to a survey of studies relevant to the validity of FIRO-B. While some were not directly aimed at the idea of validity, most involve testing predictions from FIRO theory and as such have a bearing on the construct validity of FIRO-B. Unfortunately, many of these studies utilized earlier forms of the test and therefore must be interpreted in that light.

I. FIRO-B

As noted previously, FIRO-B is the endpoint of a long evolution. Beginning with FIRO-O in 1952, Schutz continued his experimental work and was constantly refining his items until 1957 when the final set appeared. However, it was not until 1967 that the test was formally published.

A discussion of the development of FIRO-B is important for a clearer understanding of the validity research to be discussed in the following section of this chapter. Most of the studies involved earlier forms of the test and unless some description of these is given, their

relation to FIRO-B will not be evident.

FIRO-B was developed for two primary reasons:

(1) to obtain a measure of interpersonal behavior for a single individual, and, (2) "... to construct a measure that will lead to the prediction of interaction between people, based on data from the measuring instrument alone [Schutz, 1966, p. 58]." Relative to the current status of psychological measurement, the latter is an extremely ambitious goal. As Borko (1960) noted, few psychologists would expect to rely solely on the data from a single, relatively short test.

FIRO-O, a 60-item O-sort, was the initial FIRO test. The items provided a score in each of three categories: "personalness (desire for close personal relations), dependence (willingness to give up personal decision making prerogatives to authority symbols) and aggressiveness (willingness to express oneself in a group) [Schutz, 1966, p. 206]."

In an attempt to establish a more objective scoring method, FIRO-1 was created. Schutz applied the Guttman scalogram technique and established thirteen scales which gave measures of the same three variables outlined for FIRO-O.

The multiplicity of scales in FIRO-1 was troublesome especially when only three need areas were of interest. FIRO-2 was designed to overcome this difficulty and with

fewer items provided measures of four variables: personal, dependent leader, dependent follower, and participant. At that point only the control area (dependent leader and dependent follower) appeared to require a description of the superior and subordinate positions. As this lack of symmetry was not entirely satisfactory, Schutz continued his work by developing FIRO-4, which "was a logical advance over FIRO-2, although very similar overall." Then came FIRO-5B3, which measured the same six variables as the final version, FIRO-B.

During this evolutionary period, two important changes took place although the basic conceptions in FIRO theory remained relatively stable. The first was the adoption of the Guttman technique as a method of test construction. Secondly, Schutz found it necessary to expand the number of basic variables measured by the test. Parallel to this expansion in the test, FIRO theory, although retaining the same three basic need areas, also became more refined as Schutz recognized the complex nature of interpersonal behavior.

The published version of FIRO-B (see Appendix B) is an attractive, easily used four page booklet. The array of fifty-four items and their corresponding answer alternatives are conveniently arranged for response recording. The subject's responses are quickly scoreable with a well designed scoring key.

Schutz states that the questionnaire can usually be administered to a group of one hundred in about fifteen minutes and with very little instruction necessary.

II. PREVIOUS VALIDITY RESEARCH

Aside from Schutz's own work, efforts to obtain validity data on FIRO-B have been extremely limited. Not only is the number of studies small but all have been restricted to investigating what would be classified in Loevinger's model as the external component on construct validity.

Kramer (1964) obtained rank order correlations between his subjects' (psychology students) six scale scores on FIRO-B and their corresponding self-ratings on the same six variables. The students were administered FIRO-B, then were asked to discuss what they felt the test was measuring. When it was established that the subjects did not clearly understand what variables were involved, they were given a lecture on the test and FIRO theory. Lastly, the students rated themselves on the six dimensions. All resultant correlations (.33 to .83), except for express inclusion (.33), were significant at the .05 level.

Kramer's findings can only be considered mildly supportive of the validity of FIRO-B. It is conceivable that the significant correlations could have occurred because of the nature of the experimental procedure. Very

little time passed between answering the questionnaire and doing the self-ratings. It is possible that with the kind of information given in the lecture combined with a recent knowledge of the test's contents and format, the students could have given ratings quite consistent with the test responses. Had the test been administered some weeks prior to the self-ratings, the evidence for the validity of FIRO-B would have been stronger.

Gard (1964) administered FIRO-B to subjects manifesting various personality disorders including paranoids, hebephrenic schizophrenics, undifferentiated schizophrenics, obsessive-compulsive neurotics, anxiety hysterics, and depressives. All were male veterans being treated in Veteran's Administration facilities. Gard had hypothesized, from FIRO theory, that these groups would reflect different interpersonal orientations.

In general, Gard's predictions were sustained by his evidence. Unfortunately, his published results did not reflect the variety of personality types he tested. However, the schizophrenics, as predicted, obtained low scores on the inclusion scales. Likewise, the neurotics (including obsessive-compulsive neurotics) showed greater variation in scores on express affection, although the hypothesized relation for want affection did not hold. The obsessive-compulsive neurotics did not score high on express control as was expected.

Regrettably, Gard's data are incompletely reported and a detailed examination of his results is not possible. No indication of how all the groups scored on all scales is given. For example, group scores are provided for only the inclusion and affection scales. The single reference to control scales occurs in the text when Gard points out that the obsessive-compulsive groups did not score as predicted on express control. Further, on the inclusion scales, where some means are given, they allow only gross comparisons, such as schizophrenics, vs. non-schizophrenics so that it is impossible to tell how the neurotic subcategories scored on those items. The reverse is true for the affection scales, although in his discussion Gard did observe that the schizophrenic category did tend to score low on affection.

In spite of the unclear presentation of results, Gard's study does contribute to a further understanding of FIRO-B's validity. The low inclusion score for schizophrenics is in keeping with their social withdrawal tendencies. Perhaps their tendency to obtain low scores on the affection scales is an aspect of this. In any event, FIRO-B appears to be related to external non-test behavior although the specific dimensions of this non-test behavior are not clearly delineated.

The remainder of the studies to be considered are from Schutz's (1966) book. Three research approaches were used: FIRO scales were compared to measured political

attitudes, individuals differing in their tendency to conform to group pressures were compared on FIRO scales and lastly, different occupational groups were analyzed for their interpersonal orientations.

In the study on political attitudes, eighty-three Harvard undergraduates were administered FIRO-4 and FIRO-5B3 along with questionnaires designed to measure political attitudes. The political measures, developed as Guttman scales having five to seven items, were designed to correspond to each of the four FIRO scales and consequently, covered the same variables. Item content related to political events rather than general interpersonal behavior.

The "Political Individual Significance Scale" contains items such as "The President must love being famous", and "Any man who gets to be President is pretty special", which correspond to the inclusion variable. Both items emphasize being outstanding and liking to be well-known. These are characteristics of inclusion behavior as noted in a previous chapter. The other political scales follow the same general plan.

The "Political Autocrat Scale" and the "Political Abdicrat Scale" correspond to the express and want aspects of control. The "Political Personal Scale" relates to the affection variable.

A chi-square test was made between the subjects' scores on each political scale and its corresponding FIRO

variable. An examination of Schutz's (1966, p. 70-72) tables of data shows, when a one-tail test is applied, that the significant relationship was found in each comparison except for that between the "Political Personal Scale" and the affection variable. As Schutz (1966) noted, "the result of this study is very encouraging for the hypothesis that interpersonal relations orientations are significantly related to specific political attitudes" (p. 72).

A point of caution, however, should be stressed. The tests being compared were intentionally similar in content and format; therefore, this similarity accounts for some degree of correlation between them. The object of the investigators was to create measures like the FIRO scales rather than to discover what dimensions actually existed in political attitudes. In a sense, the political scales were "parallel" forms of the FIRO tests. The statistical relation between them may be more an artifact of instrument structure than an indication of the relation between two constructs. That the political scales provided an adequate external criterion for validation purposes was not really demonstrated with sufficient strength. The contributions this study can make toward establishing the validity of FIRO-B are not great.

Another study discussed by Schutz (1966) deals with the important variables related to conformity to group norms. Any situation involving the application of group

pressure to bring about conforming behavior on the part of a group member clearly involves interpersonal phenomena. In such situations there are people who change their behavior in response to group pressure and there are those who do not. Using sixty Harvard freshmen as subjects, Schutz compared the interpersonal orientations of changers and non-changers.

The subjects were randomly assigned to twelve groups of five members each. Problem situations were presented and each subject recorded his own proposed solution. The groups then discussed the issues and were asked to come to a consensus concerning problem solution. Following the discussion each subject recorded, on an objective answer sheet, the nature and extent of his opinion change.

On the basis of these data, the participants were divided into changers (those who changed on at least one problem) and non-changers. The two groups were then compared on the FIRO-1 scales. Results showed that non-changers tended to be high participators who stated that being liked was unimportant to them. Changers tended to like strict rules. Schutz concluded that the need areas seemed to be related to resistance to conformity pressures.

Some criticisms can be directed at this research as evidence for the validity of the FIRO-B scales. Firstly, the experiment involved FIRO-1, a rather distant ancestor of FIRO-B. Secondly, only three of the thirteen scales in

FIRO-1 seemed related to the conformity dimension. Of these the measure of participation was the only truly meaningful finding significant beyond the .01 level of probability. The relation between changers and liking strict rules was barely significant with a one-tailed test at the .05 level. The finding that changers did not mind being disliked was actually non-significant at just below the .10 level with a one-tailed test. Because there was nothing to suggest that these relations were predicted, it is questionable that the one-tailed test was appropriate. If a two-tailed test is applied, only the relation between participating and non-changing stands as significant.

The appropriateness of the one-tailed test depends on whether or not Schutz specified, beforehand, the directions of the expected differences on each scale. The only prediction made by Schutz is found in the following statement:

Those who do conform to the group when subjected to the social pressure should score differently in an easily explicable way on FIRO-B than those who do not conform. An experiment was performed to test this prediction [Schutz, 1966, p. 74].

Clearly, no specific predictions as to the directions of the differences between the groups is given. If the above interpretation of the data is allowed to stand, then the only significant measure being provided by the FIRO-1 scales is a measure of the amount a subject likes to participate in interpersonal behavior.

The last validation procedure to be discussed, and the one employed in the present study, involves the comparison of various occupational groups as to their interpersonal characteristics. Of the techniques dealt with so far, this appears to offer the soundest method of establishing external criteria. Most occupations have strong interpersonal elements which can be analyzed and isolated. For example, an army officer is in a control situation where he is constantly directing the behavior of others. Express control behavior is an important feature of his interpersonal orientations.

An individual's chosen occupational role is likely to be compatible with his interpersonal orientations as they are developed by his childhood experiences (Schutz, 1966, p. 66). This interpretation is consistent with the theory of occupational choice outlined by Siegleman and Peck (1960). One of their major tenants states that,

... the majority of individuals in a given vocation ... share a common personality need pattern that is different from the personality need patterns characterizing individuals in other vocational groups [p. 296].

Schutz (1966) provides information on two studies, one of which is clearly an occupational study while the other involves students from three different colleges, one of which (Harvard Business School) has occupational characteristics.

In the first study, four occupational groups were

selected: airforce officers, industrial supervisors, public school administrators, and student nurses. Why these specific groups were chosen and what score patterns were expected of them is not given. As before, no real criteria relevant to the FIRO scales were established. All groups were administered FIRO-4. Table 4 contains the score patterns as presented by Schutz (1966, p. 73).

Schutz (1966) makes only the general comment that the results seem reasonable and congruent with the stereotypes of the occupational roles (p. 73). This point is most easily accepted with respect to the high control scores of the officers and supervisors. It is not as readily self-evident that the other scores should assume the obtained pattern.

TABLE 4
SCORES ON FIRO-4 OF VARIOUS OCCUPATIONAL GROUPS

	Affection	Express Control	Want Control
Officers	Low	High	High
Supervisors	High	High	High
Teachers	High	Low	Low
Nurses	High	Low	Low

Note: Reprinted from W.C. Schutz, The interpersonal under-world. Palo Alto: Science and Behavior Books, 1966.

Because of the vagueness of the presentation many questions about the study remain unanswered. First, the external criteria problem mentioned above is further confused with Schutz's change of terminology. In the text he referred to "school administrators" and in the table to "teachers". It is unlikely that we would expect identical score patterns from administrators and teachers. The obtained pattern could probably represent teachers better than administrators. If the subjects were administrators, the low control scores would be surprising. In any event, the results are not easily interpreted.

A second problem with the data also relates to clarity of interpretation. The meaning of "high" and "low" is not defined. Does it mean, for example, that the officers scored higher on the control scales than they did on affection, or does it mean their "high's" and "low's" are relative to some external standard? No hint is provided as to what that standard might be.

A third and rather curious point should be noted. No data on the inclusion variable are given, "since it did not seem relevant in the FIRO-4 form [Schutz, 1966, p. 73]." This statement is the only explanation given for the omission. Yet, the inclusion scale was relevant in the studies on political attitudes and conformity. Indeed, the findings associated with this scale tended to be among the strongest. In addition, the same interpersonal dimensions

are supposed to be involved in all situations where interpersonal orientations are important, so that there is no sound basis for claiming that one of the variables does not apply in any one of these strongly interpersonal situations. The omission of the inclusion data therefore appears to be a serious flaw in the study. As an argument for the validity of FIRO-B's six scales, the above study has little to offer. The lack of data on the inclusion scale, the inadequacy of Schutz's discussion of the findings, and the paucity of information about relevant criteria prevent the investigation from being accepted as an adequate validity study. Taken alone, it is only mildly suggestive that some of the interpersonal variables are being measured.

In another study, three Harvard groups were given FIRO-B. One group was the entire freshman class at Harvard, another, the freshman class at Radcliffe (all females), and, the third, sixty-one students enrolled in a course at the Harvard Business School. Results showed the Business School students scoring higher than the others on express control and somewhat higher on express inclusion. Although none of the other differences were significant, the Business School subjects tended to score highest over-all.

In general, on interchange with people, that is, the desire to have extensive relations with people, they score much higher than either freshman group (total of all scores = 31.42, to 28.75 for Harvard freshmen, and 26.20 for the Radcliffe freshmen). This result seems consistent with the stereotype of the businessman [Schutz, 1966, p. 74].

A striking sex difference was noted in the express control items.

As a validity study, this investigation suffers from some of the same weaknesses as Schutz's previously discussed work. No meaningful rationale seemed to exist for selecting the student groups. Consequently there was an absence of criteria corresponding to the interpersonal dimensions. However, the finding that there was a strong sex difference on the express control scale may be meaningful in terms of the roles men and women play in western culture. The only other really significant finding was that, in general, all six subtests provide a global measure of the "desire to have extensive relations with people."

Some generalizations about the validity of FIRO-B can be drawn from the studies discussed above. None are strongly supportive of the validity of the test. The central problem seems to be one of establishing adequate criterion groups for each of the need areas. Gard's study was the only one that approached this standard but his data were reported in such a way that all the relationships could not be observed.

However, some understanding of what constructs FIRO-B appears to be measuring was gained. That the express control scale is measuring a dimension related to power and authority seems established reasonably well by both occupational studies. Gard explained his failure to find a relationship

between obsessive-compulsive neurosis and express control by suggesting that the problem of control for this group is intrapersonal rather than interpersonal.

The validity of the scales for the other variables has not been convincingly established by the studies. The Harvard student study showed that the strongest measure appeared to be a global indication of the amount of interpersonal contact an individual desires. "Participation" was the only outstanding variable related to the change-nonchange dimension in the conformity study. In Gard's investigation the fact the schizophrenics tended to score low on both inclusion and affection scales suggests that both may be indicative of socially withdrawing behavior and supports the interpretation that the scales in general reflect a measure of interpersonal interaction on the part of an individual.

FIRO-B, then, appears to be a valid indicator of the amount of interpersonal interaction a person desires and also a valid indicator of the amount of controlling behavior an individual is likely to express. The validity for the test's ability to measure all three interpersonal needs has not been adequately established.

CHAPTER IV

RESEARCH PROCEDURE

The purpose of the present study was to investigate the components of construct validity as outlined by Loevinger (1957): the substantive component, the structural component, and the external component. The following discussion of the research procedures used applies only to the latter two components. The question of the substantive component can only be answered through a rational examination of the test in conjunction with related theory. Chapter V contains a discussion of how well FIRO-B's item content reflects the theoretical constructs described in Schutz's system.

I. THE STRUCTURAL COMPONENT

As noted previously, Schutz chose the Guttman scale as the ideal model for FIRO-B. To demonstrate structural validity, a correspondence between the test's inter-item structure and the structural model must be indicated by the test responses obtained from representative members of an appropriate population. For FIRO-B the appropriate population is all people (see Postulate 1).

Thus far, Schutz has shown that the test items conform to the Guttman model for the college freshmen tested in his original cross-validation procedure. To broaden

the validity base it was considered advisable to investigate the test's scale properties on non-college subjects. Validity data would then more adequately reflect the characteristics of the population covered by the theory.

Sample

The study sample consisted of one hundred and forty-four persons ranging in age from twenty-five to eighty years. Of these, twenty-four were female and one hundred and twenty were male. None had any formal schooling beyond the twelfth grade.

The subjects were members of three occupational groups (chosen to investigate the external component): forty-eight were traffic patrolmen in the City of Edmonton Police Department, forty-eight were male and female (equal numbers) volunteers registered at the Central Volunteer Bureau, and the remaining forty-eight were salesmen from five life insurance companies.

Data Collection

All subjects were administered FIRO-B. The police and sales subjects were tested in small groups while the volunteer subjects were tested individually in their homes. No record of subjects' names was kept so that anonymity could be assured.

Each subject recorded his responses on a specially designed answer sheet (see Appendix C) rather than in the

booklet. In addition, an information sheet (see Appendix E) was provided so that the subject could supply certain biographical data such as sex, age, education level, marital status, number of clubs belonged to, and number of club executive positions held. This information was then recorded on the answer sheet by the investigator along with an identification number for each subject. An IBM 1230 Optical Scorer was used to score the answer sheets.

Statistical Analysis

To investigate the Guttman scale qualities of the test, the Goodenough technique was applied. The nature of this analysis was discussed in Chapter II.

Hypothesis

Hypothesis 1. Each of the six FIRO-B scales has a reproducibility coefficient of at least .90.

II. THE EXTERNAL COMPONENT

The major weakness of the previous validity research on FIRO-B was a failure to establish adequate criteria for judging the validity of the scales. Occupations appeared to offer one of the best means for obtaining criterion groups and was therefore adopted for the present study. Each occupation served as a criterion group for one of the three need areas.

In addition to the occupation category, information on

other behavior manifestations such as club membership and the number of club executive posts held was collected so that other external criteria could be established.

Before a description of the samples, research procedures and hypotheses is undertaken, a discussion of the rationale behind the choice of occupations is important.

Inclusion

For the inclusion area, the salesmen constituted the criterion group. Schutz (1966) stated that "... the essence of a salesman's job, ordinarily, is wide and frequent contact with people or high inclusion [p. 67]." Inclusion, then, should be the salesman's predominant interpersonal need in both its express and want aspects.

Another feature of inclusion behavior involves the joining of clubs. This represents a formalized attempt to make contact with others and it may be expected that those who belong to clubs have relatively high inclusion needs. Club membership may be considered one of the possible external criteria for the inclusion scales.

Control

The selection of an occupational category reflecting the control need was more difficult than for the inclusion area. Policemen were selected, although it may be argued that they represent only the express control aspect of the total need definition. In society they are expected to

assume a directing and aggressive role in carrying out their function. The fact that they might also have a strong interpersonal need to be controlled is not readily evident. It may be most useful to consider the police sample as a criterion group corresponding to the express control scale only.

Another indication of a high express control need might be the frequency with which a person becomes a member of the executive of an organization. It is likely that only persons seeking to control others will persist in such posts.

Perhaps a possible indication of high want control would be a tendency to avoid executive offices in clubs. A follower would be an unlikely seeker of high office.

Affection

The task of finding an occupation involving high affection needs proved extremely difficult when education level was kept low. Careers in the "helping" professions usually require considerable post high-school training, thus disqualifying all such groups from the present study.

Instead of a formal occupation, registrants at the Central Volunteer Bureau were selected as a group likely to reflect high affection needs. The purpose of the volunteer is to donate his time and energy to help some less fortunate individual. Typical activities include visiting the sick and aged in institutions or their homes, providing

transportation to medical appointments, etc., helping with educational programs for children, and generally, being "friends" to those they assist. Clearly, the affection variable is important in this area.

One problem with this group should be discussed in some detail. The vast majority of those registered at the Central Volunteer Bureau are female (over 90 per cent). What implications does this have relative to the affection need? While it is certainly true that women have more time for volunteer work, it is possible that they may also have higher affection needs. To control for this possibility it was considered advisable to test equal numbers of males and females from the volunteer group.

Samples

(1) Insurance Salesmen

A total of forty-eight male life insurance salesmen ranging in age from twenty-five to sixty-two years were tested.

(2) Policemen

A total of forty-eight traffic patrolmen from the City of Edmonton Police Department were selected. All were male and ranged in age from twenty-five to forty-three years. All had a minimum of 2.5 years experience on the force.

(3) Volunteers

Twenty-four males, aged twenty-five to eighty years,

constituting almost the entire roster of active volunteer males at the Central Volunteer Bureau, were tested. An equal number of females, aged twenty-five to sixty-nine years, were randomly selected from a list of active volunteers at the Bureau.

Statistical Analysis

All hypotheses with regard to the external component were tested by a one-way analysis of variance and differences between pairs of means were tested by the Scheffe procedure. The critical significance level chosen for this study was .05.

Hypotheses

Hypothesis 2. Salesmen will score higher on express and want inclusion than the police and volunteer subjects.

Hypothesis 3. Police will score higher on the express control than the sales and volunteer subjects.

Hypothesis 4. Volunteers will score higher on the affection scales than the sales or police subjects.

Hypothesis 5. Those who belong to one or more clubs will score higher on express inclusion than those who do not belong to any clubs.

Hypothesis 6. Those who have held or are holding executive offices in clubs will score higher on express control than club members who do not hold such offices.

Hypothesis 7. Those club members who have never held executive offices will score higher on want control than those who do hold such offices.

CHAPTER V

RESULTS AND DISCUSSION

I. THE SUBSTANTIVE COMPONENT

No empirical data were collected on the substantive component. Instead, this section contains a brief discussion of the logical relations that exist between FIRO-B and FIRO theory. Two aspects of the substantive component will be examined: firstly, how well the theory accounts for the present FIRO-B items, and secondly, how adequately the items cover the theory. The latter point is, in a sense, concerned with those items which may have been excluded from the test.

Actually, of the three components of construct validity, the substantive is the least meaningful in terms of a cross-validation analysis. This aspect seems most relevant in the early stages of test construction when the author, having a large pool of items, is faced with the problem of selecting some items while rejecting others. Once the test has been published, it is impossible to repeat or re-examine the original process of item selection from a content oriented point of view. While it is possible to generate another pool of items and then proceed with the examination, this process may not be altogether fair to the test author. It is unlikely that the second pool of items would be equivalent to the first.

However, some judgment, albeit probably weak, can be made concerning the adequacy with which an already published test meets the requirements of the substantive component. A joint examination of the test and its corresponding behavior theory is required.

An inspection of the FIRO-B items (Appendix B) shows that their relationship to the constructs described in FIRO theory are quite clear. Words such as "join", "included", and "participate" from the inclusion scales are readily related to the inclusion area. In a like manner terms such as "dominant", "take charge", and "influence" in the control items correspond to the control needs. The affection items contain such terms as "friendly", "close", and "personal". The theory, then, appears to account for the FIRO-B items and the relations between the test and theory are obvious. In fact, the items appear patently transparent, clearly indicating to the testee what variables are being measured. Surprisingly enough, however, the identity of the variables may remain obscure in spite of the apparent clarity in the item content and wording. Kramer (1964) found that psychology students could not name the need areas simply from exposure to the items.

While it is clear that each test item's presence can be explained by theory, it is not so convincing that the items adequately reflect the scope of the constructs postulated. Whole scales appear to be rewordings of only a

few ideas with scoring criteria being changed to obtain scalability. For example, the affection items are mostly variations on the idea of having "close, personal relations with people". Item 4 reads, "I try to have close relationships with people"; item 8, "I try to have close personal relationships with people"; item 21, "I try to have close personal relationships with people": item 23, "I try to get close and personal with people"; and item 12, "I try to get close and personal with people". These six very similar items constitute two-thirds of the items in the express affection scale. While some of the item wordings are identical, their meanings are slightly different by virtue of the two answer modes associated with the items. Items 4, 8 and 12 are associated with the usually-never response alternatives, while 21, 23 and 27 use the most people-nobody alternatives.

In spite of this attempt to obtain variation with different response alternatives, the variety of items does not adequately cover the affection construct as outlined in Schutz's (1966) second chapter. While some of the milder forms of affection behavior such as "like", "personal" and "cool" are reflected in the items, none seem to be measuring the more intense aspects such as "love" and "hate". The same general problem exists for the other scales in the test. They seem uncomfortably narrow relative to the behaviors supposedly associated with them.

A further problem concerning item content is the possible similarity of behaviors reflected in the inclusion and affection scales. The affection items are concerned with having relationships with others and these relations are defined on a friendly to cool continuum. The content of the inclusion items limits them to reflecting participation and being with people. While these may be logically separable concepts, it is not clear whether or not the items actually define the important and essential difference between them. In a practical situation, trying "to be with people" in the inclusion sense may be identical to trying "to have close relationships with people" in the affection sense. Perhaps items such as "I find it easy to like or even come to love people when I get to know them" and "I like to have people around me even when I don't feel close to them" would more accurately allow the separate concepts of affection and inclusion to be measured. Item wordings, then, may not actually reflect the essential distinctions between the inclusion and affection areas and therefore would not distinguish the relative strengths of those needs in people. In contrast, the item content of the control scales is quite different. The terminology is uniquely related to power concepts.

A final point that may be dealt with concerns the want and express aspects of the need areas. Item content seems to adequately reflect this distinction. An inspection of each

item shows clearly whether it belongs to a want or an express scale.

In general, there is some doubt as to the adequacy with which FIRO-B satisfies the substantive component of construct validity. While the relations between the items and the theory are fairly clear from the standpoint of content, questions arise from two other problems. Firstly, the ranges of the constructs are not reflected widely enough in the variety of published items. Much of the interpersonal behavior goes unmeasured. Secondly, there is some doubt that the item content is really able to distinguish between affection and inclusion in most practical situations.

II. THE STRUCTURAL COMPONENT

Hypothesis 1

The hypothesis that all scales would obtain a .90 reproducibility coefficient was only partially supported. Table 5 contains the reproducibility coefficient for each subtest. Only the two affection scales satisfied the requirements of a Guttman scale, although all of the others except express control were fairly close (.85 to .88).

In addition, a factor analysis of the data was carried out. The results are presented in Table 6 and suggest that two major factors exist in the items. The first includes all inclusion and affection scales with a fairly significant loading from the express control items. The second factor

TABLE 5
REPRODUCIBILITY COEFFICIENTS FOR EACH OF THE SIX
FIRO-B SCALES

Scale	Reproducibility Coefficient
Express Inclusion	.80
Want Inclusion	.86
Express Control	.85
Want Control	.88
Express Affection	.91
Want Affection	.90

TABLE 6
UNROTATED PRINCIPAL AXES FACTORS BASED ON SCALE SCORES OF
ONE HUNDRED FORTY-FOUR SUBJECTS

Scales	Factors	
	1	2
Express Inclusion	.781	.014
Want Inclusion	.810	-.006
Express Control	.392	-.659
Want Control	.145	.829
Express Affection	.797	.129
Want Affection	.744	.039

consists of the express and want control items which form a bipolar dimension with express control loading with a weight of $-.659$ and the want control loading at $.829$. A table indicating intercorrelations among the scales is given in Appendix F.

III. THE EXTERNAL COMPONENT

To evaluate the adequacy with which FIRO-B met the requirements of the external component of construct validity, three groups including volunteers, salesmen, and policemen were compared on the six FIRO-B scales. Their group mean scores and standard deviations are given in Table 7.

TABLE 7

MEAN SCORES AND STANDARD DEVIATIONS FOR THE VOLUNTEER, SALES, AND POLICE GROUPS ON THE FIRO-B SCALES

	Volunteers		Salesmen		Policemen	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Express Inclusion	4.40	2.11	5.67	2.15	3.77	2.20
Want Inclusion	3.20	3.37	4.75	3.51	2.75	3.25
Express Control	2.29	2.12	4.97	2.45	2.85	2.71
Want Control	3.58	1.67	2.73	1.61	3.17	1.93
Express Affection	3.90	1.97	5.42	2.49	3.33	2.07
Want Affection	4.96	2.50	6.19	2.68	4.73	2.22

Hypothesis 2

The hypothesis that salesmen would score higher on express and want inclusion than volunteers and policemen was partially supported (see Tables 8 and 9).

A Scheffe comparison (Table 8) among means on express inclusion shows sales subjects scoring significantly higher than either volunteers or policemen. The latter two groups did not differ significantly in their mean scores.

For the want inclusion scale, Table 9 shows the salesmen scoring highest, with volunteers ranking second, and the policemen obtaining the lowest scores. While there was no differences between the salesmen and volunteers and between volunteers and policemen, the salesmen scored significantly higher on want inclusion than the policemen. In this instance Hypothesis 2 was not fully supported.

Hypothesis 3

The hypothesis that police subjects would score higher on express control than either volunteers or salesmen was not supported. Table 10 shows that salesmen scored significantly higher than volunteers and policemen. The latter two groups did not differ from one another.

Hypothesis 4

The hypothesis that volunteers would score higher than the other two groups on express and want affection was not supported by the data (see Tables 11 and 12). In neither

TABLE 8

EXPRESS INCLUSION GROUP MEANS, ANALYSIS OF
VARIANCE, AND SCHEFFE COMPARISONS
AMONG MEANS

(a) Mean Scores on Express Inclusion

Groups	Means
Salesmen	5.67
Volunteers	4.40
Policemen	3.77

(b) Summary Analysis of Variance

Source	MS	df	F	p
Groups	44.8	2	9.62	.001
Error	4.6	141		

(c) Probability Matrix Showing Significance of Differences
Among Means Under a Scheffe Comparison

	Policemen	Volunteers	Salesmen
Policemen	1.000	.368	.001
Volunteers		1.000	.018
Salesmen			1.000

TABLE 9

WANT INCLUSION GROUP MEANS, ANALYSIS OF VARIANCE,
AND SCHEFFE COMPARISONS AMONG MEANS

(a) Mean Scores on Want Inclusion

Groups	Means
Salesmen	4.75
Volunteers	3.20
Policemen	2.75

(b) Summary Analysis of Variance

Source	MS	df	F	p
Groups	52.69	2	4.62	.011
Error	11.42	141		

(c) Probability Matrix Showing Significance of Differences Among Means Under a Scheffe Comparison

	Policemen	Volunteers	Salesmen
Policemen	1.000	.802	.017
Volunteers		1.000	.086
Salesmen			1.000

TABLE 10

EXPRESS CONTROL GROUP MEANS, ANALYSIS OF VARIANCE,
AND SCHEFFE COMPARISONS AMONG MEANS

(a) Mean Scores on Express Control

Groups	Means
Salesmen	4.97
Policemen	2.85
Volunteers	2.29

(b) Summary Analysis of Variance

Source	MS	df	F	p
Groups	96.44	2	16.17	<.001
Error	5.96	141		

(c) Probability Matrix Showing Significance of
Differences Among Means Under a Scheffe
Comparison

	Volunteers	Policemen	Salesmen
Volunteers	1.000	.531	<.001
Policemen		1.000	<.001
Salesmen			1.000

TABLE 11

EXPRESS AFFECTION GROUP MEANS, ANALYSIS OF VARIANCE,
AND SCHEFFE COMPARISONS AMONG MEANS

(a) Mean Scores on Express Affection

Groups	Means
Salesmen	5.42
Volunteers	3.90
Policemen	3.33

(b) Summary Analysis of Variance

Source	MS	df	F	p
Groups	55.76	2	11.65	.001
Error	4.75	141		

(c) Probability Matrix Showing Significance of
Differences Among Means Under a Scheffe
Comparison

	Policemen	Volunteers	Salesmen
Policemen	1.000	.454	.001
Volunteers		1.000	.004
Salesmen			1.000

TABLE 12

WANT AFFECTION GROUP MEANS, ANALYSIS OF VARIANCE,
AND SCHEFFE COMPARISONS AMONG MEANS

(a) Mean Scores on Want Affection

Groups	Means
Salesmen	6.19
Volunteers	4.96
Policemen	4.73

(b) Summary Analysis of Variance

Source	MS	df	F	p
Groups	29.52	2	4.81	.010
Error	6.13	141		

(c) Probability Matrix Showing Significance of
Differences Among Means Under a Scheffe
Comparison

	Policemen	Volunteers	Salesmen
Policemen	1.000	.902	.018
Volunteers		1.000	.056
Salesmen			1.000

case did the volunteers obtain the highest ranking scores. Instead, on express affection, the salesmen scored significantly higher than did the volunteer and police subjects who did not differ substantially from each other (Table 11).

Table 12 shows that on the want affection items the salesmen scored highest, volunteers second highest, and policemen lowest. Repeating the pattern of the want inclusion scale, the sales subjects differed significantly from the policemen but not from the volunteers. Police and volunteer subjects had almost identical means.

Hypothesis 5

The hypothesis that club members would score higher on express inclusion than those who did not belong to clubs was supported. Table 13 shows that club members obtained a mean score of 4.97 which differed significantly from the mean of 3.94 for persons not belonging to a club.

Hypothesis 6

The hypothesis that club members who held or had held executive office would score higher on express control than those who had not held executive office was not supported. Of the ninety-four subjects who indicated club membership, thirty had never held executive office and sixty-four indicated that they were or had been on club executives. An inspection of Table 14 reveals that no difference in

TABLE 13

CLUB AND NON-CLUB GROUP MEANS ON EXPRESS INCLUSION
AND SUMMARY ANALYSIS OF VARIANCE

(a) Group Means on Express Inclusion				
Groups	Means	N		
Club	4.97	94		
Non-club	3.94	50		

(b) Summary Analysis of Variance				
Source	MS	df	F	p
Groups	34.5	1	6.88	.01
Error	5.01	142		

TABLE 14

EXECUTIVE AND NON-EXECUTIVE GROUP MEANS ON EXPRESS
CONTROL AND SUMMARY ANALYSIS OF VARIANCE

(a) Group Means on Express Control				
Groups	Means	N		
Executive	3.34	64		
Non-executive	3.20	30		

(b) Summary Analysis of Variance				
Source	MS	df	F	p
Groups	.42	1	.06	.802
Error	6.67	92		

express control scores was obtained by the two groups.

Hypothesis 7

The hypothesis that club members not holding executive offices would score higher on want control than club executives was not supported (see Table 15).

TABLE 15

EXECUTIVE AND NON-EXECUTIVE GROUP MEANS ON WANT CONTROL AND SUMMARY ANALYSIS OF VARIANCE

(a) Mean Scores				
Groups	Means	N		
Non-executive	3.57	30		
Executive	3.20	64		

(b) Summary Analysis of Variance				
Source	MS	df	F	p
Groups	2.70	1	.83	.363
Error	3.24	92		

Discussion of the External Component

An inspection of the results quickly reveals that two of the six hypotheses were confirmed -- Hypotheses 2 and 5. Both were concerned with the inclusion variable thus supporting Schutz's contention that FIRO-B is measuring a variable which may be referred to an "inclusion behavior". That other variables are also being measured is not as readily apparent.

Table 6 indicates that salesmen scored highest on every scale except want control, a result in keeping with Schutz's study on the college students in which the Harvard Business School subjects scored highest overall. The same pattern of scores for the sales, volunteer, and police groups on the inclusion and affection scales suggests that they are measuring similar, if not the same, constructs. That a common dimension exists in all the items (except want control) is indicated by the consistently high scores of the sales group, the high interactors. Instead of measuring express and want inclusion, express and want affection, and express control, the items may simply be providing an indication of the magnitude of interpersonal behavior. This seems especially true of the inclusion and affection scales, and is in keeping with their similar contents as noted in the section on the substantive component.

The evidence with respect to the control scales is less clear. The fact that the salesmen ranked highest on express control and lowest on want control (see Appendix E) suggests that the control scales constitute a bipolar dimension which reflects a construct that is quite different from affection and inclusion. In the latter two scales, express and want appear to be very similar ideas so that a high score on express is accompanied by a high score on the want items. Item content here suggests no substantial distinction between express and want. For example, item 9

(Express Inclusion) "I try to include other people in my plans" is very similar in meaning to item 34 (Want Inclusion) "I like people to include me in their activities". The content of the control scales, on the other hand, sharply defines opposite ideas in the express and want scales. Item 54 (Express Control) "I take charge of things when I'm with people" is clearly the opposite of item 20 (Want Control) "I let other people take charge of things".

Schutz (1966) found, in his college study involving Radcliffe girls, that women tended to score lower on express control, thus providing evidence for a control construct. Table 16 contains data from the present study which shows that the sex difference was not replicated. Significant differences occurred between salesmen and male and female volunteers as well as the police subjects on express control. Neither volunteer group differed from the police group and the male and female volunteers did not differ from each other. The basic difference is between the sales group and all other regardless of sex. In fact, no sex differences occurred on any of the variables between any of the groups (see Appendix G).

The control scales remain a problem in that their content obviously relates to a control construct, yet the group that should have scored highest did not. Perhaps the "interpersonal interaction" dimension strongly influences the response tendency on these items so that the score on

TABLE 16

EXPRESS CONTROL MEAN SCORES FOR MALE AND FEMALE
VOLUNTEERS, SALESMEN AND POLICEMEN: ANALYSIS OF
VARIANCE AND SCHEFFE COMPARISONS AMONG MEANS

(a) Group Means on Express Control

Groups	Means	N
Salesmen	4.98	48
Male Volunteers	3.04	24
Policemen	2.85	48
Female Volunteers	1.54	24

(b) Summary Analysis of Variance

Source	MS	df	F	p
Groups	73.29	3	12.61	.001
Error	5.81	140		

(c) Probability Matrix Showing Significance of
Differences Among Means Under a Scheffe
Comparison

	Female Vol.	Policemen	Male Vol.	Salesmen
Female Vol.	1.000	.197	.205	.001
Policemen		1.000	.992	.001
Male Vol.			1.000	.019
Salesmen				1.000

this scale provides a measure of control in a highly interacting individual. This would mean that if the amount of a person's interpersonal contact were low, he would not score high on express control even though he may be a highly controlling individual. Such an explanation may be in keeping with Gard's (1963) finding that obsessive-compulsive subjects failed to score high on express control. Such people very often are socially reserved, distant, discouraging of friendly advances, and inhibited (Laughlin, 1967, p. 323). In a sense then, the express control items may be partly dependent on a "participation" factor and not really a pure measure of controlling behavior.

Evidence concerning the external component strongly indicates that FIRO-B is primarily a measure of the amount of interpersonal behavior exhibits. Supportive of this conclusion are the results of the factor analysis carried out under the structural component evaluation. A strong major factor consisting of the inclusion and affection scales, and to some extent the express control scale, was demonstrated. This factor could be interpreted as primarily an indicator of interpersonal activity. Actually, this interpretation is identical with Schutz's own conclusion with regard to his college study where he pointed out that FIRO-B provides a measure of "... the desire to have extensive relations with people [Schutz, 1966, p. 74]."

IV. THE CONSTRUCT VALIDITY OF FIRO-B

The data and discussion in the previous sections of this chapter indicate that FIRO-B is not a valid measure of all the constructs in FIRO theory. There are serious questions concerning the adequacy with which the test meets the requirements of each of the three components of construct validity as outlined by Loevinger (1957).

Under the substantive component, there is serious doubt that items sufficiently cover the scope of the theory suggesting that some potential test items were eliminated from the test without a justifiable rationale. From the standpoint of item content, the scales cannot measure the total range of interpersonal behavior as outlined in FIRO theory.

Because the inclusion and control scales failed to obtain a .90 reproducibility coefficient, none of them can be classified as Guttman scales. However, all but the express inclusion scale obtained coefficients between .85 and .88, figures very close to the .90 critical level which suggest that the items may not be so easily dismissed as not forming Guttman scales, especially as the more rigorous Goodenough technique was applied. Militating against a cautious acceptance of the scales as adequately reproducible is the fact that the items are highly homogeneous and appear to consist of rewordings of a few basic ideas (see

substantive component). Had the item content been more varied the obtained coefficients would have appeared more impressive.

In addition, a factor analysis of test structure indicated that two fairly major factors are being measured. The first, and most general factor, appears to reflect the amount of interpersonal activity a person exhibits. The second factor appears to measure a control dimension. In general, FIRO-B does not adequately meet the requirements of the structural component for the sample tested.

An evaluation of the external component suggests that all six basic constructs are not being measured. Instead, the items seem meaningfully sensitive to only two or three, these being amount of interpersonal interaction, want control and express control. Clearly, the requirement that test scores be predictably consistent with external criteria was not met in this study.

As a measuring instrument for the dimensions of Schutz's interpersonal theory, FIRO-B cannot be said to possess construct validity.

CHAPTER VI

IMPLICATIONS AND SUMMARY

I. IMPLICATIONS

The results of the present study have two important implications for further investigation concerning FIRO theory. The first relates specifically to FIRO-B. The conclusion that the test does not have adequate construct validity and that it is primarily providing a measure of the magnitude of interpersonal behavior is admittedly tentative. The data collected did not directly permit such an interpretation and it was only inferred from the score patterns. A study of a different design would allow a direct investigation of this explanation of the data. For example, it may be possible to select from discussion groups those persons judged as high interactors and those classified as low interactors. It could be hypothesized that the high interactors would obtain greater scores on the inclusion and affection scales as well as on express control. If supported, this hypothesis would substantially confirm a major conclusion of the present study.

Secondly, if the results and interpretation of this investigation are accurate, the fact that FIRO-B lacks construct validity poses a serious problem for FIRO theory as a system capable of producing hypotheses open to

scientific research. FIRO-B is the only instrument that purports to measure Schutz's interpersonal need constructs and if FIRO-B is not valid, then no means that are entirely consistent with the theoretical structure exist for operationally defining the needs. Without an operationally sound basis for defining the constructs, none of the hypotheses with regard to the interpersonal needs can be adequately investigated. Without a valid measuring instrument, FIRO theory could not move beyond the level of speculation and would add very little to scientific psychology. Before FIRO theory can develop further, more effective and meaningful measures of the need areas must be created. Items must be constructed that will distinguish among the need areas and yet remain less dependent on the actual degree on interpersonal activity in which a person participates.

II. SUMMARY

The purpose of the present study was to evaluate the construct validity of FIRO-B, the principal measuring instrument in Schutz's (1966) theory of interpersonal behavior. Schutz postulates that each individual possesses three interpersonal needs, each varying in strength in accordance with childhood interpersonal experiences. The needs cover three areas: inclusion, control, and affection. FIRO-B was designed to measure each of the three needs, both in their express and want aspects, with six 9-item

Guttman scales.

The research procedure was based on Loevinger's (1957) theory of psychometrics. The central concept of her theory is construct validity, which consists of three necessary components: the substantive component, the structural component, and the external component.

FIRO-B was evaluated as only partially meeting the requirements of the substantive component because the test items did not appear to adequately reflect the breadth of the constructs of FIRO theory.

Results showed that FIRO-B also failed to fully satisfy the requirements of the structural component. The structural model chosen for the test was the Guttman scale, and to show that FIRO-B had structural validity it was necessary to demonstrate that each of the six FIRO scales would obtain at least the minimum reproducibility coefficient of .90 required of the Guttman structure. To evaluate this component, the test was administered to one hundred forty-four non-college adults ranging in age from twenty-five to eighty years. The Goodenough technique was applied to the scale analysis which showed that the inclusion and control scales failed to obtain high enough reproducibility coefficient, so that only the affection scales, with coefficients of .90 and .91 could be classified as Guttman scales. However, it should be noted that want inclusion and the control scales attained

coefficients between .85 and .88, levels quite respectable for psychological instruments. On the other hand, as noted above, the item content was constricted and limited to only a few ideas so that the magnitude of the coefficients was not as significant as it would have been had item content been more varied.

To evaluate the external component, three occupations were chosen as criterion groups, each corresponding to one of the three need areas. Salesmen represented the inclusion variable, policemen the control variable, and service volunteers were hypothesized to be representative of the affection variable. In addition, other behavioral data was collected. Forty-eight subjects were tested in each group with the volunteer group consisting of males and females in equal numbers. The same one hundred forty-four subjects were involved in the analysis of the structural component discussed above.

Only two of the six hypotheses relevant to the external component were supported. Both involved the inclusion scales and the results were in keeping with the evidence from the previous validity research, which found measures of the inclusion area more valid than the others. The general pattern of scores suggested that two, or perhaps three, constructs, rather than six, were being measured by the FIRO-B items. The major construct reflected in the test appeared to be predominantly inclusion-like (interpersonally

participating) behavior with the result that the FIRO scales seem to provide a measure of the amount of interpersonal interaction a person customarily seeks. The control items, as well as reflecting the "participation" factor, appeared to relate to a control dimension of a bipolar nature consisting of the express and want aspects of controlling behavior.

The results have important implications for the research productivity of FIRO theory and suggest that a new set of items be developed to provide a more valid measure of the three need areas: inclusion, control, and affection.

R E F E R E N C E S

REFERENCES

- American Psychological Association, Committee on Test Standards, APA, AERA-NCME. Standards for educational and psychological tests and manuals. Washington, D.C., APA, 1966.
- Benne, K. and Sheats, P. Functional roles of group members. Journal of Social Issues, 1948, 4, 41-49.
- Borko, H. Review of W.C. Schutz, FIRO-A three-dimensional theory of interpersonal behavior. Educational and Psychological Measurement, 1960, 28, 210-211.
- Campbell, D.T. and Fiske, D.W. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 1959, 56, 81-105.
- Champney, H. The variables of parent behavior. Journal of Abnormal and Social Psychology, 1951, 36, 525-542.
- Cronbach, L.J. and Meehl, P.E. Construct validity in psychological tests. Psychological Bulletin, 1955, 52, 281-302.
- Douglas, J.D. Review of W.C. Schutz, FIRO-A three-dimensional theory of interpersonal behavior. American Journal of Sociology, 1961, 66, 543.
- Edwards, A.L. On Guttman's scale analysis. Educational and Psychological Measurement, 1948, 8, 313-318.
- Edwards, A.L. Techniques of attitude scale construction. New York: Appleton-Century-Croft, 1957.
- Festinger, L. The treatment of qualitative data by "scale analysis". Psychological Bulletin, 1947, 44, 149-161.
- Freud, S. Libidinal types, in Collected papers. London: Hogarth, 1950.
- Fromm, E. Man for himself. New York: Rinehart, 1947.
- Gard, J.G. Interpersonal orientations in clinical groups. Journal of Abnormal and Social Psychology, 1964, 69, 516-521.
- Goodenough, W.H. A technique for scale analysis. Educational and Psychological Measurement, 1944, 4, 179-190.

- Gulliksen, H. Theory of mental tests. New York: Wiley, 1950.
- Guttman, L. The basis for scalogram analysis, in S.A. Stouffer et al., Measurement and prediction. Princeton: Princeton University Press, 1950. Pp. 60-90.
- Horney, K. Our inner conflicts. New York: Norton, 1945.
- Kramer, E. A contribution toward the validation of the FIRO-B questionnaire. Journal of Projective Techniques and Personality Assessment, 1967, 31, 80-81.
- Laughlin, H.P. The neuroses. Washington: Butterworth, 1967.
- Leary, T. The interpersonal diagnosis of personality. New York: Ronald, 1957.
- Loevinger, J. Objective tests as instruments of psychological theory. Psychological Reports, 1957, (Monograph Supplement 9).
- Schutz, W.C. FIRO-A three-dimensional theory of interpersonal behavior. New York: Holt, Rinehart and Winston, 1958. (Republished: The interpersonal underworld. Palo Alto: Science and Behavior Books, 1966.
- Siegleman, M. and Peck, R. Personality patterns related to occupational roles. Genetic Psy. Monograph, 1960, 61, 291-349.
- Tukey, J.W. Discussion: Symposium on statistics for the clinician. Journal of Clinical Psychology, 1950, 6, 61-74.
- White, B.W. and Saltz, E. The measurement of reproducibility. Psychological Bulletin, 1957, 54, 81-99.

A P P E N D I C E S

A P P E N D I X A

FIRO Theory

FIRO: A THREE-DIMENSIONAL THEORY OF INTERPERSONAL BEHAVIOR

Below are the propositions expressing what are taken to be the most fundamental ideas embodied in the theory. In the ideal instance, all theorems in the system are derivable from these postulates. In the present instance, they represent the best estimate of the most basic ideas in the theory presented (Schutz, 1966, pp. 196-200).

Postulate 1. THE POSTULATE OF INTERPERSONAL NEEDS

- (a) Every individual has three interpersonal needs: inclusion, control, and affection.
- (b) Inclusion, control, and affection constitute a sufficient set of areas of interpersonal behavior for the prediction and explanation of interpersonal phenomena.

Theorem 1-1. If a representative battery of measures of interpersonal behavior is factor-analyzed, the resulting factors will reasonably fall into the three need areas, inclusion, control, and affection.

Postulate 2. THE POSTULATE OF RELATIONAL CONTINUITY

An individual's expressed interpersonal behavior will be similar to the behavior he experienced in his earliest interpersonal relations, usually with his parents, in the following way:

Principle of Constancy. When he perceives his adult position in an interpersonal situation to be similar to his own position in his parent-child relation, his adult behavior positively covaries with his childhood behavior toward his parents (or significant others).

Principle of Identification. When he perceives his adult position in an interpersonal situation to be similar to his parent's position in his parent-child relation, his adult behavior positively covaries with the behavior of his parents (or significant others) toward him when he was a child.

Postulate 3. THE POSTULATE OF COMPATIBILITY

If the compatibility of one group, h , is greater than that of another group, m , then the goal achievement of h will exceed that of m .

Theorem 3-1. If the compatibility of one dyad, y , is greater than the compatibility of another dyad, x , then the members of y are more likely to prefer each other for continued personal interchange.

Theorem 3-2. If the compatibility of one group, h , is greater than the compatibility of another group, m , then the productivity goal achievement of h will exceed that of m .

Theorem 3-3. If the compatibility of one group, h ,

is greater than the compatibility of another group, m, the h will be more cohesive than m.

Theorems 3-4, 3-5, and 3-6 were not confirmed in further testing; therefore they are not included here (Schutz, 1966, p. 198).

Theorem 3-7. In the compatible groups, those predicted to be focal persons and those predicted to be main supporting members should rank one another very high on the relation "work well with".

Theorem 3-8. In all the groups focal persons will be chosen as group leaders by the group members.

Theorem 3-9. The effect of compatibility on productivity (the C-P effect) increases as the task situation requires more interchange in the three need areas.

Postulate 4. THE POSTULATE OF GROUP DEVELOPMENT

The formation and development of two or more people into an interpersonal relation (that is, a group) always proceed in the same sequence, as follows:

Principle of Group Integration. For the time period starting with the group's beginning until three intervals before the group's termination the predominant area of interpersonal behavior begins with inclusion, is followed by control, and finally by affection. This cycle may recur.

Principle of Group Resolution. The last three intervals prior to a group's anticipated termination follow the opposite sequence in that the predominant area of interpersonal behavior is first affection, then control, and finally inclusion.

Theorem 4-1. (1) If a group is high on inclusion compatibility, but low on control compatibility and affection compatibility, group members will be most compatible during the initial stages of their relation.

(2) If a group is high on control compatibility, but low on inclusion compatibility and affection compatibility, group members will be most compatible during the middle stages of their relation.

(3) If a group is high on affection compatibility, but low on inclusion compatibility and control compatibility, group members will be most compatible during the last stages of their relation prior to the separation sequence.

A P P E N D I X B

FIRO-B Questionnaire

For each statement below, decide which of the following answers best applies to you. Place the number of the answer in the box at the left of the statement. Please be as honest as you can.

1. usually 2. often 3. sometimes 4. occasionally 5. rarely 6. never

- | | |
|---|---|
| ☐ 1. I try to be with people. | ☐ 9. I try to include other people in my plans. |
| ☐ 2. I let other people decide what to do. | ☐ 10. I let other people control my actions. |
| ☐ 3. I join social groups. | ☐ 11. I try to have people around me. |
| ☐ 4. I try to have close relationships with people. | ☐ 12. I try to get close and personal with people. |
| ☐ 5. I tend to join social organizations when I have an opportunity. | ☐ 13. When people are doing things together I tend to join them. |
| ☐ 6. I let other people strongly influence my actions. | ☐ 14. I am easily led by people. |
| ☐ 7. I try to be included in informal social activities. | ☐ 15. I try to avoid being alone. |
| ☐ 8. I try to have close, personal relationships with people. | ☐ 16. I try to participate in group activities. |

For each of the next group of statements, choose one of the following answers:

1. most people 2. many people 3. some people 4. a few people 5. one or two people 6. nobody

- | | |
|--|---|
| ☐ 17. I try to be friendly to people. | ☐ 23. I try to get close and personal with people. |
| ☐ 18. I let other people decide what to do. | ☐ 24. I let other people control my actions. |
| ☐ 19. My personal relations with people are cool and distant. | ☐ 25. I act cool and distant with people. |
| ☐ 20. I let other people take charge of things. | ☐ 26. I am easily led by people. |
| ☐ 21. I try to have close relationships with people. | ☐ 27. I try to have close, personal relationships with people. |
| ☐ 22. I let other people strongly influence my actions. | |

For each of the next group of statements, choose one of the following answers:

1. most people 2. many people 3. some people 4. a few people 5. one or two people 6. nobody

- | | |
|--|---|
| ☐ 28. I like people to invite me to things. | ☐ 35. I like people to act cool and distant toward me. |
| ☐ 29. I like people to act close and personal with me. | ☐ 36. I try to have other people do things the way I want them done. |
| ☐ 30. I try to influence strongly other people's actions. | ☐ 37. I like people to ask me to participate in their discussions. |
| ☐ 31. I like people to invite me to join in their activities. | ☐ 38. I like people to act friendly toward me. |
| ☐ 32. I like people to act close toward me. | ☐ 39. I like people to invite me to participate in their activities. |
| ☐ 33. I try to take charge of things when I am with people. | ☐ 40. I like people to act distant toward me. |
| ☐ 34. I like people to include me in their activities. | |

For each of the next group of statements, choose one of the following answers:

1. usually 2. often 3. sometimes 4. occasionally 5. rarely 6. never

- | | |
|---|---|
| ☐ 41. I try to be the dominant person when I am with people. | ☐ 48. I like people to include me in their activities. |
| ☐ 42. I like people to invite me to things. | ☐ 49. I like people to act close and personal with me. |
| ☐ 43. I like people to act close toward me. | ☐ 50. I try to take charge of things when I'm with people. |
| ☐ 44. I try to have other people do things I want done. | ☐ 51. I like people to invite me to participate in their activities. |
| ☐ 45. I like people to invite me to join their activities. | ☐ 52. I like people to act distant toward me. |
| ☐ 46. I like people to act cool and distant toward me. | ☐ 53. I try to have other people do things the way I want them done. |
| ☐ 47. I try to influence strongly other people's actions. | ☐ 54. I take charge of things when I'm with people. |

A P P E N D I X C

Answer Sheet

NAME

Last

First

Middle

GRADE

Years

or

GRADE

Male

Female

DATE

Day

Month

Year

FACULTY

or

SCHOOL

Indicate response by placing a mark between the guidelines as shown in the example. Use HB pencil. Don't make marks longer than guidelines.

Example

I. D. NUMBER

0123456789

0123456789

0123456789

0123456789

0123456789

1] 1 2 3 4 5 6

2] 1 2 3 4 5 6

3] 1 2 3 4 5 6

4] 1 2 3 4 5 6

5] 1 2 3 4 5 6

6] 1 2 3 4 5 6

7] 1 2 3 4 5 6

8] 1 2 3 4 5 6

9] 1 2 3 4 5 6

10] 1 2 3 4 5 6

11] 1 2 3 4 5 6

12] 1 2 3 4 5 6

13] 1 2 3 4 5 6

14] 1 2 3 4 5 6

15] 1 2 3 4 5 6

16] 1 2 3 4 5 6

17] 1 2 3 4 5 6

18] 1 2 3 4 5 6

19] 1 2 3 4 5 6

20] 1 2 3 4 5 6

21] 1 2 3 4 5 6

22] 1 2 3 4 5 6

23] 1 2 3 4 5 6

24] 1 2 3 4 5 6

25] 1 2 3 4 5 6

26] 1 2 3 4 5 6

27] 1 2 3 4 5 6

28] 1 2 3 4 5 6

29] 1 2 3 4 5 6

30] 1 2 3 4 5 6

31] 1 2 3 4 5 6

32] 1 2 3 4 5 6

33] 1 2 3 4 5 6

34] 1 2 3 4 5 6

35] 1 2 3 4 5 6

36] 1 2 3 4 5 6

37] 1 2 3 4 5 6

38] 1 2 3 4 5 6

39] 1 2 3 4 5 6

40] 1 2 3 4 5 6

41] 1 2 3 4 5 6

42] 1 2 3 4 5 6

43] 1 2 3 4 5 6

44] 1 2 3 4 5 6

45] 1 2 3 4 5 6

46] 1 2 3 4 5 6

47] 1 2 3 4 5 6

48] 1 2 3 4 5 6

49] 1 2 3 4 5 6

50] 1 2 3 4 5 6

51] 1 2 3 4 5 6

52] 1 2 3 4 5 6

53] 1 2 3 4 5 6

54] 1 2 3 4 5 6

CAUTION - AVOID PLACING ANY MARKS AMONG THE BLACK TIMING LINES

OVERPRINT ANSWER SHEET

A P P E N D I X D

Scale Items and Keyed Response Alternatives

FIRO-B SCORING KEY

Scale items and the response alternatives which obtain a score of 1.

Express Inclusion

<u>Item</u>	<u>Keyed Alternatives</u>
1	1-2-3
3	1-2-3-4
5	1-2-3-4
7	1-2-3
9	1-2
11	1-2
13	1-2
15	1
16	1

Want Inclusion

<u>Item</u>	<u>Keyed Alternatives</u>
28	1-2
31	1-2
34	1-2
42	1-2
45	1-2
37	1
39	1
48	1-2
51	1-2

Express Control

<u>Item</u>	<u>Keyed Alternatives</u>
30	1-2-3
33	1-2-3
41	1-2-3-4
44	1-2-3
47	1-2-3
36	1-2
50	1-2
53	1-2
54	1-2

Want Control

<u>Item</u>	<u>Keyed Alternatives</u>
2	1-2-3-4
6	1-2-3-4
18	1-2-3
20	1-2-3
22	1-2-3-4
10	1-2-3
14	1-2-3
24	1-2-3
26	1-2-3

Express Affection

<u>Item</u>	<u>Keyed Alternatives</u>
4	1-2
8	1-2
17	1-2
19	4-5-6
21	1-2
12	1
23	1-2
25	4-5-6
27	1-2

Want Affection

<u>Item</u>	<u>Keyed Alternatives</u>
29	1-2
32	1-2
43	1
46	5-6
35	5-6
38	1-2
40	5-6
49	1-2
52	5-6

A P P E N D I X E

Information Sheet

PART I: BIOGRAPHICAL INFORMATION

1. Sex (encircle): Male Female
2. Level of education: _____ . Age: _____
3. Marital status (circle one): Single Married
 Separated Divorced
4. What clubs or organizations do you belong to?
 - a) _____
 - b) _____
 - c) _____
 - d) _____
 - e) _____
5. Are you a member of the executive of any club or organization? (circle one): Yes No

Could you please indicate what executive positions you are now holding in a club or organization, and those you have held in the past. For example, are you or were you ever a chairman, a president, a treasurer, a secretary, etc. in an organization.

- a) _____
- b) _____
- c) _____
- d) _____

PART II: INSTRUCTIONS FOR COMPLETING THE QUESTIONNAIRE

1. Read the instructions on the test booklet, but do not mark your answer in the booklet. An answer sheet has been provided for this purpose.
2. In making your answer, place a mark between the set of broken lines on the answer sheet corresponding to the answer number you have chosen in the test booklet. For example, if a person chose answer number "4" in response to item number 55, he would mark his answer like this:

55. =1= =2= =3= ~~=4=~~ =5= =6=

A P P E N D I X F

FIRO-B Scale Intercorrelations

INTERCORRELATIONS AMONG FIRO-B SCALES

	Express Inclusion	Want Inclusion	Express Control	Want Control	Express Affection	Want Affection
Express Inclusion	1.000	.556	.256	.111	.516	.395
Want Inclusion		1.000	.259	.094	.503	.497
Express Control			1.000	-.147	.167	.151
Want Control				1.000	.123	.044
Express Affection					1.000	.543
Want Affection						1.000

A P P E N D I X G

Volunteer, Sales, and Police Group Scores
on Want Control, Analysis of Variance and
Scheffe Comparisons Among Means

WANT CONTROL SCORES, ANALYSIS OF VARIANCE, AND
SCHEFFE COMPARISON AMONG MEANS

(i) Group Means on Want Control

Groups	Means
Volunteers	3.58
Policemen	3.17
Salesmen	2.73

(ii) Summary Analysis of Variance

Source	MS	df	F	p
Groups	8.76	2	2.89	.059
Error	3.01	141		

(iii) Probability Matrix Showing Significance of Differences Among Means Under a Scheffe Comparison

	Salesmen	Policemen	Volunteers
Salesmen	1.000	.471	.059
Policemen		1.000	.505
Volunteers			1.000

A P P E N D I X H

Scale Means and Scheffe Comparisons Among
Means on the Six FIRO-B Scales for Female
Volunteers, Male Volunteers, Salesmen, and
Policemen

GROUP MEANS FOR FEMALE VOLUNTEERS, MALE VOLUNTEERS,
SALESMEN, AND POLICEMEN: SCHEFFE COMPARISONS AMONG
MEANS FOR THE SIX FIRO-B SCALES

(a) Express Inclusion

(i) Group Means

Groups	Means
Salesmen	5.67
Female Volunteers	4.54
Male Volunteers	4.25
Policemen	3.77

(ii) Probability Matrix Showing Significance of
Differences Between Means Under a Scheffe
Comparison

	Police- men	Male Vol.	Female Vol.	Sales- men
Policemen	1.000	.835	.568	.001
Male Vol.		1.000	.975	.081
Female Vol.			1.000	.233
Salesmen				1.000

(b) Want Inclusion

(i) Group Means

Groups	Means
Salesmen	4.75
Male Volunteers	3.63
Female Volunteers	2.79
Policemen	2.75

(ii) Probability Matrix Showing Significance of
Differences Between Means Under a Scheffe
Comparison

	Police- men	Female Vol.	Male Vol.	Sales- men
Policemen	1.000	1.000	.784	.043
Female Vol.		1.000	.886	.152
Male Vol.			1.000	.623
Salesmen				1.000

(c) Express Control

(i) Group Means

Groups	Means
Salesmen	4.98
Male Volunteers	3.04
Policemen	2.85
Female Volunteers	1.54

(ii) Probability Matrix Showing Significance of Differences Between Means Under a Scheffe Comparison

	Female Vol.	Police- men	Male Vol.	Sales- men
Female Vol.	1.000	.197	.205	<.001
Policemen		1.000	.992	<.001
Male Vol.			1.000	.019
Salesmen				1.000

(d) Want Control

(i) Group Means

Groups	Means
Male Volunteers	3.79
Female Volunteers	3.38
Policemen	3.17
Salesmen	2.73

(ii) Probability Matrix Showing Significance of Differences Between Means Under a Scheffe Comparison

	Sales- men	Police- men	Female Vol.	Male Vol.
Salesmen	1.000	.680	.535	.120
Policemen		1.000	.973	.563
Female Vol.			1.000	.877
Male Vol.				1.000

(e) Express Affection

(i) Group Means

Groups	Means
Salesmen	5.42
Female Volunteers	3.99
Male Volunteers	3.83
Policemen	3.33

(ii) Probability Matrix Showing Significance of Differences Between Means Under a Scheffe Comparison

	Police-men	Male Vol.	Female Vol.	Sales-men
Policemen	1.000	.842	.730	<.001
Male Vol.		1.000	.998	.044
Female Vol.			1.000	.075
Salesmen				1.000

(f) Want Affection

(i) Group Means

Groups	Means
Salesmen	6.19
Female Volunteers	5.04
Male Volunteers	4.88
Policemen	4.73

(ii) Probability Matrix Showing Significance of Differences Among Means Under a Scheffe Comparison

	Police-men	Male Vol.	Female Vol.	Sales-men
Policemen	1.000	.997	.969	.045
Male Vol.		1.000	.997	.221
Female Vol.			1.000	.338
Salesmen				1.000

B29926